

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 16:04:45 ; Search time 2069.62 Seconds  
(without alignments)  
9782.198 Million cell updates/sec

Title: US-09-863-475A-3  
Perfect score: 1500  
Sequence: 1 CCTCCCTTAGACTCTC.....CTTACACTATTTCTTAACCA 1500

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1447.4	96.5	3607	11	AK004753	AK004753 Mus muscu
2	687.4	45.8	845	10	BI688333	BI688333 603316655
3	648	43.2	810	10	BI657849	BI657849 603284656
4	589.4	39.3	1094	10	BI674143	BI674143 602334132
5	588.2	39.2	662	10	BI694566	BI694566 602348371
6	517.8	34.5	642	10	BI6976005	BI6976005 602846032
7	491.4	32.8	666	10	BI158143	BI158143 602920783
8	488	32.5	570	10	BI653667	BI653667 603300323
9	473.4	31.6	768	10	BI646945	BI646945 603278705
10	465.4	31.0	964	10	BF539781	BF539781 602049993
11	431.2	28.7	457	9	AW762229	AW762229 ur55c07.y
12	423.6	28.2	712	10	BG914492	BG914492 602813234
13	406.4	27.1	440	9	BB862353	BB862353 BB862353
14	380	25.3	520	10	BF046412	BF046412 BP250011A
15	378	25.2	574	10	BF658605	BF658605 maa98e08.
16	353.2	23.5	1370	11	AK015047	AK015047 Mus muscu
17	315.4	21.0	493	9	AW465562	AW465562 BP230019B

c	18	309	20.6	523	9	AV603266	AV603266
	19	283.2	18.9	1207	11	AK016765	AK016765
	20	280.4	18.7	440	9	AA175441	AA175441 ms87d04.r
	21	259	17.3	536	10	BF042060	BF042060 BP250023A
	22	253.8	16.9	1229	11	AK016248	AK016248 Mus muscu
c	23	251.4	16.8	399	9	AV666881	AV666881
	24	248	16.5	383	9	BB842104	BB842104 BB842104
	25	213.8	14.3	544	9	AV666882	AV666882
	26	210	14.0	479	9	AW463741	AW463741 BP230013B
	27	209.2	13.9	475	9	AW464433	AW464433 BP230015B
c	28	199	13.3	516	10	BI399614	BI399614 MI-P-AV1-
c	29	193.8	12.9	1037	12	CNS040JM	AL268987 Tetraodon
	30	189.2	12.6	362	10	BF042944	BF042944 BP250002A
	31	176.8	11.8	530	10	BF039182	BF039182 BP250008B
	32	174.2	11.6	244	9	BB574519	BB574519 BB574519
	33	169.2	11.3	333	10	BF040713	BF040713 BP250002A
c	34	164	10.9	674	12	BH078853	BH078853 RPCI-24-3
	35	140.8	9.4	743	12	AZ899051	AZ899051 RPCI-24-2
	36	139.6	9.3	552	9	AV613965	AV613965
	37	138.6	9.2	483	10	BF544312	BF544312 UI-R-BT0-
	38	137	9.1	316	10	BM434393	BM434393 1RT09A01
	39	135.6	9.0	672	9	AL631525	AL631525 AL631525
c	40	132	8.8	500	12	AZ015258	AZ015258 RPCI-23-2
	41	131.8	8.8	562	9	BE234506	BE234506 141635 MA
	42	130.4	8.7	568	9	AA088048	AA088048 mo01607.r
c	43	128.8	8.6	628	12	AQ996632	AQ996632 RPCI-23-3
	44	128.4	8.6	544	10	BE808178	BE808178 213324 MA
	45	127	8.5	664	10	BI858477	BI858477 603391708

## ALIGNMENTS

RESULT 1

AK004753

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AK004753 3607 bp mRNA linear HTC 19-JAN-2002  
Mus musculus adult male lung cDNA, RIKEN full-length enriched  
library, clone:1200014C04:glycoprotein galactosyltransferase alpha  
1, 3, full insert sequence.  
AK004753 GI:12836166  
HTC; CAP trapper.  
Mus musculus (strain:C57BL/6J) adult male lung cDNA to mRNA,  
clone:1200014C04.  
Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (sites)  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
10349636  
2 (sites)  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159  
3 (sites)  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishikawa, K., Katsunai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujii, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system -384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)



```
QY 865 TCCGGATGCTGTGTCACCTGAACCCCTCTACATTCCTTACAAAGTCTTTGAGATCAGG 924
Db 1084 TCCGGATGCTGTGTCACCTGAACCCCTCTACATTCCTTACAAAGTCTTTGAGATCAGG 1143
QY 925 TCTGAGAAGAGTGGCAGGATATCAGCATGCGCATGAAGACCATTTGGGGAGCACATC 984
Db 1144 TCTGAGAAGAGTGGCAGGATATCAGCATGCGCATGAAGACCATTTGGGGAGCACATC 1203
QY 985 CTGCGCCACATCAGCAGGAGTGCATTCCTCTTCTGTCATGACGCTGGGATCAAGTCTTT 1044
Db 1204 CTGCGCCACATCAGCAGGAGTGCATTCCTCTTCTGTCATGACGCTGGGATCAAGTCTTT 1263
QY 1045 CAAGACAACTTCGGGTGGAACTCTGGGCCAGCTGGTAGCACAGCTCCAGGCTGGTG 1104
Db 1264 CAAGACAACTTCGGGTGGAACTCTGGGCCAGCTGGTAGCACAGCTCCAGGCTGGTG 1323
QY 1105 TACAAGGCCAGTCCCGAGAGTTTCACTATGAGAGCGGGAACCTGTCGGCCGCTACATT 1164
Db 1324 TACAAGGCCAGTCCCGAGAGTTTCACTATGAGAGCGGGAACCTGTCGGCCGCTACATT 1383
QY 1165 CCATTCCGAGAGGGGATTTTACTACACGCGGCCATTTTGGAGGAAGCGCTACTCAC 1224
Db 1384 CCATTCCGAGAGGGGATTTTACTACACGCGGCCATTTTGGAGGAAGCGCTACTCAC 1443
QY 1225 ATTCTCAACTCACCAGGAGTCTTTAAGGGATCTCTCAGGACAAAGAACATGACATA 1284
Db 1444 ATTCTCAACTCACCAGGAGTCTTTAAGGGATCTCTCAGGACAAAGAACATGACATA 1503
QY 1285 GAAGCCAGTGGCATGATGAGAGCCACTCAACAAATACCTTCTTTTCAACAACCCACT 1344
Db 1504 GAAGCCAGTGGCATGATGAGAGCCACTCAACAAATACCTTCTTTTCAACAACCCACT 1563
QY 1345 AAAATCCTATCTCCAGAGTATTCCTGGGACTATCAGATAGGCGCTTCAGATATTTAA 1404
Db 1564 AAAATCCTATCTCCAGAGTATTCCTGGGACTATCAGATAGGCGCTTCAGATATTTAA 1623
QY 1405 AGTGTCAAGTAGCTGGCAGACAAAGAGTATAATTTGGTTAGAATAATGTGACTT 1464
Db 1624 AGTGTCAAGTAGCTGGCAGACAAAGAGTATAATTTGGTTAGAATAATGTGACTT 1683
QY 1465 CAAATTGTGATGGAATCTGACACTATTCT 1495
Db 1684 CAAATTGTGATGGAATCTGACACTATTCT 1714

RESULT 2
LOCUS BI688333
DEFINITION 845 bp mRNA linear EST 18-SEP-2001
603316655F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5356621 5',
mRNA sequence.
ACCESSION BI688333
VERSION BI688333.1 GI:15650962
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 845)
NTH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1906 row: n column: 14
High quality sequence stop: 704.
Location/Qualifiers
```

## source

```
1. .845
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5356621"
/clone_lib="NCI_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally; Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT 222 a 233 c 212 g 178 t
ORIGIN
```

```
Query Match 45.8%; Score 687.4; DB 10; Length 845;
Best Local Similarity 98.9%; Pred. No. 1.1e-186;
Matches 713; Conservative 0; Mismatches 6; Indels 2; Caps 2;

QY 37 CGATTCTGCTGAAGACCTCGCGCTCTCAGGCTCTGGGAGTT-GGAACCCCTGTACCTTTCCT 95
Db 1 CGATTCTGCTGAAGACCTCGCGCTCTCAGGCTCTGGGAGTGGGAACCCCTGTACCTTTCCT 60
QY 96 TTCTCTGCTGAGCCCTGCTCTCTTAGGCGAGGCGAGAGCTCGACAGAACTCGGTTCGCTTT 155
Db 61 TTCTCTGCTGAGCCCTGCTCTCTTAGGCGAGGCGAGAGCTCGACAGAACTCGGTTCGCTTT 120
QY 156 GCTGTTTGTGTTGGAGGGAACACAGCTGAGGATGAGGCTGACTTTGAACTCAAGATCT 215
Db 121 GCTGTTTGTGTTGGAGGGAACACAGCTGAGGATGAGGCTGACTTTGAACTCAAGATCT 180
QY 216 GCTTACCCCACTCTCTTGGAAATTAAGGCGCTGTACTACATTTGCTCGGACCTTAAGATTTT 275
Db 181 GCTTACCCCACTCTCTTGGAAATTAAGGCGCTGTACTACATTTGCTCGGAGCTTAAGATTTT 240
QY 276 CATGATCACTATCTTCAAGATCTCCATGTCAACAAGATCTCCATGTCAAGATCCAAAGTC 335
Db 241 CATGATCACTATCTTCAAGATCTCCATGTCAACAAGATCTCCATGTCAAGATCCAAAGTC 300
QY 336 AGAACAAGTCTTCCATCTCCATCTCAAGATCTGATCAGAGGAGAGAGAAATATGATGTCAGGG 395
Db 301 AGAACAAGTCTTCCATCTCCATCTCAAGATCTGATCAGAGGAGAGAGAAATATGATGTCAGGG 360
QY 396 AAAAGTAATCTCTGTGATGCTGATTTCAACCGCTGTTGCTGCTGTTTGGGAATATGT 455
Db 361 AAAAGTAATCTCTGTGATGCTGATTTCAACCGCTGTTGCTGCTGTTTGGGAATATGT 420
QY 456 CAACAGAAATTTCCAGAGGTTGGTGAGAACAGATGGCAGAGGAGCTGGTGGTCCCAAGCTG 515
Db 421 CAACAGAAATTTCCAGAGGTTGGTGAGAACAGATGGCAGAGGAGCTGGTGGTCCCAAGCTG 480
QY 516 GTTTAAATATGGAGCCACACAGTATCAAGAGACACAGCTAGAGGAGCGGAGAGAGAAAGGG 575
Db 481 GTTTAAATATGGAGCCACACAGTATCAAGAGACACAGCTAGAGGAGCGGAGAGAGAAAGGG 540
QY 576 TAGAAATGGAGATCGCATTTCAAGAGCCTCAGCTATGGAGCTGGTTCAATCCAAAGAACCG 635
Db 541 TAGAAATGGAGATCGCATTTCAAGAGCCTCAGCTATGGAGCTGGTTCAATCCAAAGAACCG 600
QY 636 CCCGGATGTTTTCACAGTGTACCCCGTGGAGGCGCCGATTTGTGTGGAGGAGCACTTATGA 695
Db 601 CCCGGATGTTTTCACAGTGTACCCCGTGGAGGCGCCGATTTGTGTGGAGGAGCACTTATGA 660
QY 696 CACAGCTCTCTGGAA-AAGTACTACGCCACACAGAAACTCACTGTGGGCTGACAGTGT 754
Db 661 CACAGCTCTCTGGAAACAAAGTACTACGCCACACAGAAACTCACTGTGGGCTGACAGCGG 720
QY 755 T 755
Db 721 T 721
```

QY	308	GAAGAAGGATAGAAATGGAGATCGCATTTGAAGAGCCCTCAGCTATGAGGACCTGGTTCATATCCA	627
Db	271	GAAGAAGGATAGAAATGGAGATCGCATTTGAAGAGCCCTCAGCTATGAGGACCTGGTTCATATCCA	330
QY	628	AGAAGCCGCCCGATGTTTGTACAGTGACCCCGTGAAGGCGCCGATTTGTGTGGGAGGCC	687
Db	331	AGAAGCCGCCCGATGTTTGTACAGTGACCCCGTGAAGGCGCCGATTTGTGTGGGAGGCC	390
QY	688	ACTTATGACACAGCTCTGCTGGAAAAGTACTACGCCACACACAACTCCTGTGGGGCTG	747
Db	391	ACTTATGACACAGCTCTGCTGGAAAAGTACTACGCCACACACAACTCCTGTGGGGCTG	450
QY	748	ACAGTCTTTTCTGCTGGAAAAGTACTACGCCACACACAACTCCTGTGGGGCTGCTCT	807
Db	451	ACAGTCTTTTCTGCTGGAAAAGTACTACGCCACACACAACTCCTGTGGGGCTGCTCT	510
QY	808	GACATGCTACTTTCATGTTTGGCCATCGGGTGCATATTTTACGTCATGATAGACGACCTCC	867
Db	511	GACATGCTACTTTCATGTTTGGCCATCGGGTGCATATTTTACGTCATGATAGACGACCTCC	570
QY	868	CGGATGCCCTGCTGTCGACCTGAACCTCTACATTCCTTACAAAGCTTTTGAAGCTCAGGTCT	927
Db	571	CGGATGCCCTGCTGTCGACCTGAACCTCTACATTCCTTACAAAGCTTTTGAAGCTCAGGTCT	630
QY	928	GAGAAGAGCTGGCAGGATAT-CAGCATGATGGCATGAAGACCA-TTGGGAGACACATCC	985
Db	631	GAGAAGAGCTGGCAGGATAT-CAGCATGATGGCATGAAGACCA-TTGGGAGACACATCC	690
QY	986	TGGCCACATCCAGCAGGAGTGA-CATTCTCTTTTCATGCGAGGTGGATCAAGTCTTTT	1044
Db	691	TGGCCACATCCAGCAGGAGTGA-CATTCTCTTTTCATGCGAGGTGGATCAAGTCTTTT	750
QY	1045	CAAGCAACTT-CGGGGTGGAAACTCTGGGCGACCTGCTGACC	1085
Db	751	CCAGCAACTTCCCGGCTGGAAACTCTGGGCGACCTGCTGACC	792
RESULT	4		
LOCUS	BG174143	1094 bp	linear
DEFINITION	602334132F1 NCI_CCAP_Mam1 Mus musculus cDNA clone IMAGE:4457334 5',		
ACCESSION	BG174143		
VERSION	BG174143.1	GI:12680846	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 1094)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-r@mail.nih.gov		
	Tissue Procurement: Gilbert Smith, Ph.D.		
	cDNA Library Preparation: Life Technologies, Inc.		
	DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	Plate: LLAM10253 row: p column: 07		
	High quality sequence stop: 662.		
FEATURES	Location/Qualifiers		
SOURCE	1..1094		
	/organism="Mus musculus"		
	/strain="FVB/N"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:4457334"		
	/clone_lib="NCI_CCAP_Mam1"		
	/tissue_type="tumor, biops sample"		





```

Db 538 TGAGATCAGGTCTGAGAGAGGTGGCAGGATATCAGCATGATGCGCATGAAGACCATTTGG 597
Qy 975 GGAGCACATCTGCGCCACATCCAGCAGCGAGTGGACTTCTCTC-TTCTGCGATGGACGTGG 1033
Db 598 GGAGCACATCTGCGCCACATCCAGCAGCGAGTGGACTTCTCTC-TTCTGCGATGGACGTGG 657
Qy 1034 ATCAA 1038
Db 658 ATCAA 662

RESULT 6
Bg976005 602846032F1 NCI_CGAP_Mam6 642 bp mRNA linear EST 12-JUN-2001
LOCUS 602846032F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4976863 5',
DEFINITION mRNA sequence.
ACCESSION Bg976005
VERSION Bg976005.1 GI:14363642
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryote; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus.
1 (bases 1 to 642)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10970 row: o column: 08
High quality sequence start: 3
High quality sequence stop: 642.
FEATURES
source
1. .642
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4976863"
/clone_lib="NCI_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT 160 a 147 c 199 g 136 t
ORIGIN
Query Match 34.5%; Score 517.8; DB 10; Length 642;
Best Local Similarity 97.3%; Pred. No. 7e-138;
Matches 571; Conservative 0; Mismatches 7; Indels 9; Gaps 4;

Qy 543 AGAAGACACGTAGAGGACGAGAGAGAAAGGGGTAGAAATCGAGATCGCATTCGAAGAGCC 602
Db 60 AGAAGACACGTAGAGGACGAGAGAGAAAGGGGTAGAAATCGAGATCGCATTCGAAGAGCC 119
Qy 603 TCAGCTATGGACTGGTTCAATCCAAAGAACCGCCCGGATGTTTTCAGATGACGCCCGGTG 662
Db 120 TCAGCTATGGACTGGTTCAATCCAAAGAACCGCCCGGATGTTTTCAGATGACGCCCGGTG 179
Qy 663 GAAGGCCCGCATTTGTGTGGGAAGGCACTTATGACACAGCTCTGCTGGAAGAAAGTACTACGC 722
Db 180 GAAGGCCCGCATTTGTGTGGGAAGGCACTTATGACACAGCTCTGCTGGAAGAAAGTACTACGC 239

```

```

Qy 723 CACACAGAACTCACTGTGGGGCTGACAGTG-TTTTGCTGTGGAAAGTACATTTGAGCATT 781
Db 240 CACACAGAACTCACTGTGGGGCTGACAGTGTTTGTGCTGTGGAAAGTACATTTGAGCATT 299
Qy 782 ACTTAGAAGACTTTCTGTGAGTCTGCTGACATGTACTTTCATGTGTGGCCATCGGGTCTATAT 841
Db 300 ACTTAGAAGACTTTCTGTGAGTCTGCTGACATGTACTTTCATGTGTGGCCATCGGGTCTATAT 359
Qy 842 TTTACGTCAATGATA--GAGACACCTCCGGATGCGCTGCTGCTGACCTGAACCTCTAC 898
Db 360 TTTACGTCAATGATAAGACACCTCCCGGATGCGCTGCTGCTGACCTGAACCTCTAC 419
Qy 899 ATTCTTTACAAGTCTTTGAGATCAGGTCTGAGAAAGAGTGGCAGGATATCAGCATGATGC 958
Db 420 ATTCTTTACAAGTCTTTGAGATCAGGTCTGAGAAAGAGTGGCAGGATATCAGCATGATGC 479
Qy 959 GCATGAAGACCATTTGGGAGGACACATCTTGGCCACATCCAGCAGGAGTGCAGTTCCTCT 1018
Db 480 GCATGAAGACCATTTGGGAGGACACATCTTGGCCACATCCAGCAGGAGTGCAGTTCCTCT 539
Qy 1019 TCTGCATGGAGCTGGATCAAGTCTTTCAAGACAACCTTCGGGGTGAACCTCTGGCCAGC 1078
Db 540 TCTGCATGGAGCTGGATCAAGTCTTTCAAGACAACCTTCGGGGTGAACCTCTGGCCAGC 596
Qy 1079 TGTAGCACAGCTCCAGGCCCTGGT-GGTACAAGCCAGTCCCGAGAA 1124
Db 597 -TGTAGCACAGCTCCAGGCCCTGGTGGGTACAAGGACAGTCCCGAGAA 642

RESULT 7
Bg158143 602920783F1 NIH_CGAP_Mam3 666 bp mRNA linear EST 05-JUL-2001
LOCUS 602920783F1 NIH_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5061116 5',
DEFINITION mRNA sequence.
ACCESSION Bg158143
VERSION Bg158143.1 GI:14618144
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryote; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus.
1 (bases 1 to 666)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1165 row: e column: 21
High quality sequence stop: 665.
FEATURES
source
1. .666
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:5061116"
/clone_lib="NIH_CGAP_Mam3"
/tissue_type="tumor, gross tissue"
/lab_host="DH10B"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert 2 kb. Library constructed by Life
Technologies, catalog #12017-018. Investigators providing
samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference
for transgenic model: Xu et al., Nature Genetics 22, 37-43
(1999). Note: this is a NCI_CGAP Library."
BASE COUNT 185 a 158 c 160 g 163 t

```

ORIGIN

Query Match 32.8%; Score 491.4; DB 10; Length 666;  
Best Local Similarity 90.8%; Pred. No. 2.9e-130;  
Matches 595; Conservative 0; Mismatches 16; Indels 44; Gaps 5;  
QY 36 CCGATTCTGCTGAAGACCTCGCGCTCTCAGGCTCTGGAGTTGGAACCCCTGTACCTTCCT 95  
DB 1 CCGATTCTGCTGAAGACCTCGCGCTCTCAGGCTCTGGAGTTGGAACCCCTGTACCTTCCT 60  
QY 96 TTCTCTCTGCTGACCCCTGCTCTTAGGAGGCCAGCTCGACAGCTCGGCTGCTGCTTT 155  
DB 61 TTCTCTCTGCTGACCCCTGCTCTTAGGAGGCCAGCTCGACAGCTCGGCTGCTGCTTT 120  
QY 156 GCTGTTTGTCTTGGAGGAACACAGCTGAGGATGAGGCTGACTTTGAACCTCAAGAGATCT 215  
DB 121 GCTGTTTGTCTTGGAGGAACACAGCTGAGGATGAGGCTGACTTTGAACCTCAAGAGATCT 180  
QY 216 GCTTACCCCTGCTCTGGAATTAAGGCTGCTGACTACATTTGCTGCTGCTGCTGCTGCT 275  
DB 181 GCTTACCCCTGCTCTGGAATTAAGGCTGCTGACTACATTTGCTGCTGCTGCTGCTGCT 240  
QY 276 CATGATCACTATGCTTCAAGATCTCCATGTCACACAGATCTCCATGTCACAGATCCAACTC 335  
DB 241 CATGATCACTATGCTTCAAGATCTCCATGTCACACAGATCTCCATGTCACAGATCCAACTC 300  
QY 336 AGAAACAAGTCTTCCATCTCCTCAAGATCTGGATCACAGGAGAAATTAATGAATGCTCAAGGG 395  
DB 301 AGAAACAAGTCTTCCATCTCCTCAAGATCTGGATCACAGGAGAAATTAATGAATGCTCAAGGG 360  
QY 396 AAAAGTAATCTGCTGATGCTGATGCTCAACCGTGGTGTGCTGCTGCTGCTGCTGCTGCT 455  
DB 361 AAAAGTAATCTGCTGATGCTGATGCTCAACCGTGGTGTGCTGCTGCTGCTGCTGCTGCT 420  
QY 456 CAACAG-----AATTCACAGAGTTGGTGA 479  
DB 421 CAACAGCCCAAGAGGCTCTTCTTGTGGATATATACACAAAATTCACAGAGTTGGTGA 480  
QY 480 GAACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 535  
DB 481 GAACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
QY 536 GTTATCAAGAGCAACGTAAGAGGACGGA---GAGAAAGGTTAGAAATGAGATGCTG 591  
DB 541 GTTATCAAGAGCAACGTAAGAGGACGGA---GAGAAAGGTTAGAAATGAGATGCTG 600  
QY 592 ATTGAAGAGCTCAGCTATGGGACTGTTCAATCCAAAGAACCCCGGCTGTTT 646  
DB 601 ATTGAAGAGCTCAGCTATGGGACTGTTCAATCCAAAGAACCCCGGCTGTTT 655

RESULT 8

BI653667 570 bp mRNA linear EST 12-SEP-2001  
LOCUS 603300333F1 NTH\_CGAP\_Mam3 Mus musculus cDNA clone IMAGE:5340868 5',  
DEFINITION mRNA sequence.  
ACCESSION BI653667  
VERSION BI653667.1 GI:15567903  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 570)  
NIH-MGC <http://mgc.nhl.nih.gov/>.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLML1865 row: n column: 05  
High quality sequence stop: 570.  
FEATURES  
Location/Qualifiers  
1..570  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5340868"  
/clone\_lib="NTH\_CGAP\_Mam3"  
/tissue\_type="tumor, gross tissue"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert 2 kb. Library constructed by Life  
Technologies, catalog #12017-018. Investigators providing  
samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference  
for transgenic model: Xu et al., Nature Genetics 22, 37-43  
(1999). Note: this is a NCI\_CGAP Library."  
BASE COUNT 162 a. 115 c 156 g 137 t  
ORIGIN  
Query Match 32.5%; Score 488; DB 10; Length 570;  
Best Local Similarity 93.7%; Pred. No. 2.5e-129;  
Matches 534; Conservative 0; Mismatches 0; Indels 36; Gaps 1;  
QY 362 CTGGATCACAGGAGAAATAAATGAATGTCAGGGAGAAAGTAATCCTGTGATGCTGATTC 421  
DB 1 CTGGATCACAGGAGAAATAAATGAATGTCAGGGAGAAAGTAATCCTGTGATGCTGATTC 60  
QY 422 TCTCAACCGTGGTGTGCTGCTTTTGGGAATATGTCACACAG----- 461  
DB 61 TCTCAACCGTGGTGTGCTGCTTTTGGGAATATGTCACACAGCCCAAGAGGCTCTTCTTGT 120  
QY 462 -----AATTCACAGAGTTGGTGTGAGAACAGATGGCAGAGGACTGGTGGT 505  
DB 121 GGATATATCACAAAATTCACAGAGTTGGTGTGAGAACAGATGGCAGAGGACTGGTGGT 180  
QY 506 TCCCAAGCTGGTTTAAATATGGGACCCACAGTTATCAAGAAGACACAGTGAAGAGCGGA 565  
DB 181 TCCCAAGCTGGTTTAAATATGGGACCCACAGTTATCAAGAAGACACAGTGAAGAGCGGA 240  
QY 566 GAGAAAGGTTAGAAATGAGATGCGATTGAAGAGCTCAGCTATGGAGCTGGTTCATC 625  
DB 241 GAGAAAGGTTAGAAATGAGATGCGATTGAAGAGCTCAGCTATGGAGCTGGTTCATC 300  
QY 626 CAAAGAACCCCGGATGTTTTCACAGTGAACCCCGTGGAAAGCCGATTTGTGTGGGAAG 685  
DB 301 CAAAGAACCCCGGATGTTTTCACAGTGAACCCCGTGGAAAGCCGATTTGTGTGGGAAG 360  
QY 686 GCACTTATGACAGCTGCTGCTGGAAGTACTACGCCACACAGAACTCCTGTTGGGGC 745  
DB 361 GCACTTATGACAGCTGCTGCTGGAAGTACTACGCCACACAGAACTCCTGTTGGGGC 420  
QY 746 TGACAGTGTGCTGTGGGAAAGTACATTGAGCAATTTACTTAGAAGACTTCTTGAGTCTG 805  
DB 421 TGACAGTGTGCTGTGGGAAAGTACATTGAGCAATTTACTTAGAAGACTTCTTGAGTCTG 480  
QY 806 CTGACATGTACTTTCATGTTGGCCATCGGCTCATATTTTACGTGATGATAGACACCT 865  
DB 481 CTGACATGTACTTTCATGTTGGCCATCGGCTCATATTTTACGTGATGATAGACACCT 540  
QY 866 CCCGGATGCTGCTGCTGACCTGAAACCTC 895  
DB 541 CCCGGATGCTGCTGCTGACCTGAAACCTC.570  
RESULT 9  
BI646945  
LOCUS BI646945 768 bp mRNA linear EST 12-SEP-2001

Db	487	ATGCTGATTTGCTCAACCGTGGTTTGTGCTGTGTTTGGGAATATGTCAACAGCCCAAGAGGCT	546
Qy	445	-----TGGGAATATGCTCAACAGAAATTCACAGAGGTTGGTGAGAACAGATGGCAGAAGG	496
Db	547	TTTTCTTGTGGATATATCACACAAAATTCAGAGGTTGGTGAGAACAGATGGCAGAAGG	606
Qy	497	ACTGGTGGTTCCTCCCAAGCTGGTTTAAAAATGGGACCCACAGTTATCAAGAGACAAACGTPAG	556
Db	607	ACTGGTGGTTCCTCCCAAGCTGGTTTAAAAATGGGACCCACAGTTATCAAGAGACAAACGTPAG	666
Qy	557	AA-CGACGGAGAGAAAAGGTTAGAAATGGAGATCGCATTTGAAGAGCCCTCAGCTATGGGAC	615
Db	667	AAGGAGCGGAGAGAAAAGGTTAGAAATGGAGATCGCATTTGAAGAGCCCTCAGCTATGGGAC	726
Qy	616	TGGTTCAATCCAAAGAACCGCCCGGATGTTT	646
Db	727	TGGTTCAATCCAAAGAACCGCCCGGATGTTT	757
RESULT	10		
LOCUS	BF539781	964 bp mRNA linear	EST 11-DEC-2000
DEFINITION	602049993F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4189174 5',		
ACCESSION	BF539781	mRNA sequence.	
VERSION	BF539781.1	GI:11627162	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 964)		
JOURNAL	NTF-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgapbs-f@mail.nih.gov">cgapbs-f@mail.nih.gov</a> Tissue Procurement: Jeffrey E. Green, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLAM9514 row: j column: 23 High quality sequence stop: 607. Location/Qualifiers 1. 964 /organism="Mus musculus" /strain="FVB/N" /db_xref="taxon:10090" /clone="IMAGE:4189174" /clone_lib="NCI_CGAP_SG2" /lab_host="DH10B (T1 phage-resistant)" /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site:1; Note1: Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."		
FEATURES	source		
BASE COUNT	267 a 237 c 252 g 208 t		
ORIGIN			
Query Match	31.0%;	Score 465.4;	DB 10; Length 964;
Best Local Similarity	79.2%;	Pred. No. 1.1e-122;	
Matches 672; Conservative	0;	Mismatches 101;	Indels 76; Gaps 7;
Qy	33	GTACCGATTCTGCTGAAGACCTCGCGCTCTCAGGCTCTGGAGTTGGAACCCCTGTACCTT	92
Db	1	GTACCGATTCTGCTGAAGACCTCGCGCTCTCAGGCTCTGGAGTTGGAACCCCTGTACCTT	60
Qy	93	CTTTTCCTCTGCTGAGCCCTGCCTCTCTTAGGCAGGCCAGAGCTCGACAGAACTCGGTTC	152
Db	61	CTTTTCCTCTGCTGAGCCCTGCCTCTCTTAGGCAGGCCAGAGCTCGACAGAACTCGGTTC	120

QY 153 TTTGCTGTTTCTTGGAGGAAACACAGCTGACGATGAGGCTGACTTTGAACCTCAAGAGA 212  
 Db 121 T-----GCTGTGTTCTGGAGGAAACACAGCTGACGATGAGGCTGAC-TTGAACCTCAAGAGA 175  
 QY 213 TCTGCTTACCCAGTCTCTCTGGAAATTAAGGCTGTACTACTATCTTCCCTGGACCTAAGAT 272  
 Db 176 TCTGCTTACCCAGTCTCTCTGGAAATTAAGGCTGTACTACTATCTTCCCTGGACCTAAGAT 235  
 QY 273 TTTTCATGATCAGCTATGCTTCAAGATCTCCATGTCACAAAGATCTCCATGTCAGATGCCAA 332  
 Db 236 TTTTCATGATCAGCTATGCTTCAAGATCTCCATGTCACAAAGATCTCCATGTCAGATGCCAA 295  
 QY 333 GTCAGAAACAGTCTTCCATCTCAAGATCTGGATCACAGGAGAAAATAATGAATGTCAA 392  
 Db 296 GTCAGAAACAGTCTTCCATCTCAAGATCTGGATCACAGGAGAAAATAATGAATGTCAA 355  
 QY 393 GGGAAAAGTAATCTGTGTGATGCTGATGCTCAACCGTGGTGTCTGTTTGGGAATA 452  
 Db 356 GGGAAAAGTAATCTGTGTGATGCTGATGCTCAACCGTGGTGTCTGTTTGGGAATA 415  
 QY 453 TGTCAACAGAAATCCAGAGGTTGGTGAGAAACAGATGGCAGAGGACTGGTGGTTCCCAAG 512  
 Db 416 TGTCAACA----- 423  
 QY 513 CTGCTTTAAAAATGGACCCACAGTATTATCAAGAAGACAAAGCTAGAGGAGGAGAGAA- 571  
 Db 424 -----GCACCCACAGTTTATCAAGAAGACAAAGCTAGAGGAGGAGAGAAAC 469  
 QY 572 AGGGTAGAATGAGATCGCATTTGAAGAGCCTCAGCTATGGGACTGTTCAATCCAAAGA 631  
 Db 470 AGGGTAGAATGAGATCGCATTTGAAGAGCCTCAGCTATGGGACTGTTCAATCCAAAGA 529  
 QY 632 ACCGCCCGGATGTTTGAACAGTCAACCCCGTGGAGGCGCGGATGTTGTGGGAAGGCACCT 691  
 Db 530 ACCGCC--GGATGTTTGAACAGTCAACCCCGTGGAGGCGCGGA-TGTGTGGGAAGGCACCT 586  
 QY 692 ATGACACAGCTGCTGTGGAAAGTACTACGCCACACAGAAACTCAGTGTGGGCTGACAG 751  
 Db 587 ATGACACAGTCTGCTGGAAAGTCTTACGCCACACAGAAACTCAGTGTGGGCGCTGAC 646  
 QY 752 TGTCTGCTGGGAAGTACATTTAGCATTACATTTAGAACACTTCTTGGAGTCTGCTGACA 811  
 Db 647 AAGGGTGTCTGGGGGAACGACCTGGGACCTTAATTTAAACACTCGGGGAACCTGCGGACAG 706  
 QY 812 TGTACTTCAATG-CTTGGCCATCGGTCATATTTACGTCATGATAGACAGACACCTCCCGG 870  
 Db 707 TGTCTCCGGTGTGTGGGGCAAGGGGCCANTACACGACGTGATAGACAAAACCCCGG 766  
 QY 871 ATGCTGTCT 879  
 Db 767 GAGTAAGTC 775

RESULT 11  
 AW762229  
 LOCUS  
 DEFINITION ur55c07.y1 NCI\_CGAP\_Mam3 Mus musculus cDNA clone IMAGE:3154188 5' similar to gb:M26925 Mouse galactosyltransferase mRNA, complete cds (MOUSE); mRNA sequence.

ACCESSION AW762229  
 VERSION AW762229.1 GI:7694150  
 SOURCE house mouse.  
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 457)  
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Other\_ESTs: ur55c07.x1

Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs@mail.nih.gov  
 Tissue procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 image.llnl.gov/image/html/iresources.shtml

MGI:1056944  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 399.  
 Location/Qualifiers  
 1. .457  
 /organism="Mus musculus"  
 /strain="129 - C57/B6 - FVBN"  
 /db\_xref="taxon:10090"  
 /clone\_image="IMAGE:3154188"  
 /clone\_lib="NCI\_CGAP\_Mam3"  
 /tissue\_type="tumor, gross tissue"  
 /dev\_stage="10 months"  
 /lab\_host="DH10B"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: Salt; Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999)."  
 22, 37-43 (1999)."

BASE COUNT 112 a 116 c 106 g 123 t  
 ORIGIN

Query Match 28.7%; Score 431.2; DB 9; Length 457;  
 Best Local Similarity 99.3%; Pred. No. 5.2e-113;  
 Matches 433; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 26 ATGAGAAGTACCGATTCTGCTGAAGACCTCGCGCTCTCAGGCTCTGGAGCTTGGAAACCTT 85  
 Db 22 ATCAGAAGTACCGATTCTGCTGAAGACCTCGCGCTCTCAGGCTCTGGAGCTTGGAAACCTT 81  
 QY 86 GTACCTTCTCTCTGCTGAGCCCTGCTCTTAGGAGGCGCAGAGCTCGACAGAACT 145  
 Db 82 GTACCTTCTCTCTGCTGAGCCCTGCTCTTAGGAGGCGCAGAGCTCGACAGAACT 141  
 QY 146 CGTTTCTTGTGCTTGTGAGGGAACACAGCTGAGGCTGAGGCTGAGCTTGAACCT 205  
 Db 142 CGTTTCTTGTGCTTGTGAGGGAACACAGCTGAGGCTGAGGCTGAGCTTGAACCT 201  
 QY 206 CAAGAGATCTGCTTACCCAGTCTCCTGGAAATTAAGGCCCTGTACTACATTTGCCCTGGAC 265  
 Db 202 CAAGAGATCTGCTTACCCAGTCTCCTGGAAATTAAGGCCCTGTACTACATTTGCCCTGGAC 261  
 QY 266 CTAAGATTTTCATGATCAGTACTATGCTTCAAGATCTCCATGTCACAAAGATCTCCATGTCAA 325  
 Db 262 CTAAGATTTTCATGATCAGTACTATGCTTCAAGATCTCCATGTCACAAAGATCTCCATGTCAA 321  
 QY 326 GATCCAAAGTCAGAAACAAGCTTCCATCCCTCAAGATCTGGATCACAGGAGAAAATAATGA 385  
 Db 322 GATCCAAAGTCAGAAACAAGCTTCCATCCCTCAAGATCTGGATCACAGGAGAAAATAATGA 381  
 QY 386 ATGTCAAGGGAAGTAATCCTTGTGATGCTGATGCTTCAACCGTGGTGTGTGCTGTGTTTT 445  
 Db 382 ATGTCAAGGGAAGTAATCCTTGTGATGCTGATGCTTCAACCGTGGTGTGTGCTGTGTTTT 441  
 QY 446 GGAATATGTCAACAG 461  
 Db 442 GGAATATGTCAACAG 457

RESULT 12  
 BG914492  
 LOCUS

BG914492 712 bp mRNA linear EST 05-JUN-2001









GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 17:50:51 ; Search time 3258.53 seconds  
(without alignments)  
9959.199 Million cell updates/sec

Title: US-09-863-475A-3  
Perfect score: 1500  
Sequence: 1 CCTTCCTGTGAGACTTTC.....CTTGACACTATTCTTAACCA 1500

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_NA\_Main:\*

1: /cgn2\_6/ptodata/2/pna/PCTUS\_COMB.seq:\*

2: /cgn2\_6/ptodata/2/pna/US06\_COMB.seq:\*

3: /cgn2\_6/ptodata/2/pna/US07\_COMB.seq:\*

4: /cgn2\_6/ptodata/2/pna/US080\_COMB.seq:\*

5: /cgn2\_6/ptodata/2/pna/US081\_COMB.seq:\*

6: /cgn2\_6/ptodata/2/pna/US082\_COMB.seq:\*

7: /cgn2\_6/ptodata/2/pna/US083\_COMB.seq:\*

8: /cgn2\_6/ptodata/2/pna/US084\_COMB.seq:\*

9: /cgn2\_6/ptodata/2/pna/US085\_COMB.seq:\*

10: /cgn2\_6/ptodata/2/pna/US086\_COMB.seq:\*

11: /cgn2\_6/ptodata/2/pna/US087\_COMB.seq:\*

12: /cgn2\_6/ptodata/2/pna/US088\_COMB.seq:\*

13: /cgn2\_6/ptodata/2/pna/US089\_COMB.seq:\*

14: /cgn2\_6/ptodata/2/pna/US090\_COMB.seq:\*

15: /cgn2\_6/ptodata/2/pna/US091\_COMB.seq:\*

16: /cgn2\_6/ptodata/2/pna/US092\_COMB.seq:\*

17: /cgn2\_6/ptodata/2/pna/US093\_COMB.seq:\*

18: /cgn2\_6/ptodata/2/pna/US094\_COMB.seq:\*

19: /cgn2\_6/ptodata/2/pna/US095A\_COMB.seq:\*

20: /cgn2\_6/ptodata/2/pna/US095B\_COMB.seq:\*

21: /cgn2\_6/ptodata/2/pna/US095C\_COMB.seq:\*

22: /cgn2\_6/ptodata/2/pna/US095D\_COMB.seq:\*

23: /cgn2\_6/ptodata/2/pna/US096A\_COMB.seq:\*

24: /cgn2\_6/ptodata/2/pna/US096B\_COMB.seq:\*

25: /cgn2\_6/ptodata/2/pna/US096C\_COMB.seq:\*

26: /cgn2\_6/ptodata/2/pna/US096D\_COMB.seq:\*

27: /cgn2\_6/ptodata/2/pna/US096E\_COMB.seq:\*

28: /cgn2\_6/ptodata/2/pna/US097A\_COMB.seq:\*

29: /cgn2\_6/ptodata/2/pna/US097B\_COMB.seq:\*

30: /cgn2\_6/ptodata/2/pna/US097C\_COMB.seq:\*

31: /cgn2\_6/ptodata/2/pna/US098A\_COMB.seq:\*

32: /cgn2\_6/ptodata/2/pna/US098B\_COMB.seq:\*

33: /cgn2\_6/ptodata/2/pna/US098C\_COMB.seq:\*

34: /cgn2\_6/ptodata/2/pna/US099A\_COMB.seq:\*

35: /cgn2\_6/ptodata/2/pna/US099B\_COMB.seq:\*

36: /cgn2\_6/ptodata/2/pna/US099C\_COMB.seq:\*

37: /cgn2\_6/ptodata/2/pna/US100\_COMB.seq:\*

38: /cgn2\_6/ptodata/2/pna/US101\_COMB.seq:\*

39: /cgn2\_6/ptodata/2/pna/US6000\_COMB.seq:\*

40: /cgn2\_6/ptodata/2/pna/US6001\_COMB.seq:\*

41: /cgn2\_6/ptodata/2/pna/US6002\_COMB.seq:\*

42: /cgn2\_6/ptodata/2/pna/US6003\_COMB.seq:\*

43: /cgn2\_6/ptodata/2/pna/US6004\_COMB.seq:\*

44: /cgn2\_6/ptodata/2/pna/US6005\_COMB.seq:\*

45: /cgn2\_6/ptodata/2/pna/US6006\_COMB.seq:\*

46: /cgn2\_6/ptodata/2/pna/US6007\_COMB.seq:\*

47: /cgn2\_6/ptodata/2/pna/US6008\_COMB.seq:\*

48: /cgn2\_6/ptodata/2/pna/US6009\_COMB.seq:\*

49: /cgn2\_6/ptodata/2/pna/US6010\_COMB.seq:\*

50: /cgn2\_6/ptodata/2/pna/US6011\_COMB.seq:\*

51: /cgn2\_6/ptodata/2/pna/US6012\_COMB.seq:\*

52: /cgn2\_6/ptodata/2/pna/US6013\_COMB.seq:\*

53: /cgn2\_6/ptodata/2/pna/US6014\_COMB.seq:\*

54: /cgn2\_6/ptodata/2/pna/US6015\_COMB.seq:\*

55: /cgn2\_6/ptodata/2/pna/US6016\_COMB.seq:\*

56: /cgn2\_6/ptodata/2/pna/US6017\_COMB.seq:\*

57: /cgn2\_6/ptodata/2/pna/US6018\_COMB.seq:\*

58: /cgn2\_6/ptodata/2/pna/US6019\_COMB.seq:\*

59: /cgn2\_6/ptodata/2/pna/US6020\_COMB.seq:\*

60: /cgn2\_6/ptodata/2/pna/US6021\_COMB.seq:\*

61: /cgn2\_6/ptodata/2/pna/US6022\_COMB.seq:\*

62: /cgn2\_6/ptodata/2/pna/US6023\_COMB.seq:\*

63: /cgn2\_6/ptodata/2/pna/US6024\_COMB.seq:\*

64: /cgn2\_6/ptodata/2/pna/US6025\_COMB.seq:\*

65: /cgn2\_6/ptodata/2/pna/US6026\_COMB.seq:\*

66: /cgn2\_6/ptodata/2/pna/US6027\_COMB.seq:\*

67: /cgn2\_6/ptodata/2/pna/US6028\_COMB.seq:\*

68: /cgn2\_6/ptodata/2/pna/US6029\_COMB.seq:\*

69: /cgn2\_6/ptodata/2/pna/US6030\_COMB.seq:\*

70: /cgn2\_6/ptodata/2/pna/US6031\_COMB.seq:\*

71: /cgn2\_6/ptodata/2/pna/US6032\_COMB.seq:\*

72: /cgn2\_6/ptodata/2/pna/US6033\_COMB.seq:\*

73: /cgn2\_6/ptodata/2/pna/US6034\_COMB.seq:\*

74: /cgn2\_6/ptodata/2/pna/US6035\_COMB.seq:\*

75: /cgn2\_6/ptodata/2/pna/US6036\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1500	100.0	1500	7	US-08-379-040-1
2	1500	100.0	1500	12	US-08-823-489-3
3	1500	100.0	1500	22	US-09-593-316-9
4	1500	100.0	1500	33	US-09-863-475-3
5	1500	100.0	1500	33	US-09-863-475A-3
6	1500	100.0	1500	35	US-09-946-034-1
7	1447.4	96.5	3455	75	US-60-360-207-12027
8	1408.4	93.9	3450	5	US-08-188-607B-9
9	1408.4	93.9	3450	7	US-08-378-617-9
10	1408.4	93.9	3450	13	US-08-984-900-9
11	1005.8	67.1	1131	6	US-08-213-200A-3
12	1005.8	67.1	1131	15	US-09-173-270-3
13	726.6	48.4	1423	6	US-08-214-580-2
14	726.6	48.4	1423	6	US-08-260-201-5
15	726.6	48.4	1423	6	US-08-278-282-5
16	726.6	48.4	1423	15	US-09-137-517-2
17	705	47.0	1091	16	US-09-230-091-3
18	702.8	46.9	1140	15	US-09-173-270-1
19	693.8	46.3	1131	6	US-08-213-200A-1
20	693.8	46.3	1131	22	US-09-593-316-5
21	693.8	46.3	1131	36	US-09-994-427A-14
22	693.8	46.3	1131	36	US-09-995-419A-8
23	692.2	46.1	1131	36	US-09-994-427A-16
24	684.2	45.6	1269	6	US-08-328-933A-1
25	684.2	45.6	1269	13	US-08-929-940-1
26	684.2	45.6	1269	22	US-09-593-316-7
27	684.2	45.6	1412	13	US-08-984-900-7
28	679.4	45.3	1412	5	US-08-188-607B-7
29	679.4	45.3	1412	7	US-08-378-617-7
30	663.6	44.2	1127	16	US-09-230-091-1
31	645.6	43.0	1303	22	US-09-593-316-1
					Sequence 1, Appl
					Sequence 3, Appl
					Sequence 9, Appl
					Sequence 3, Appl
					Sequence 3, Appl
					Sequence 3, Appl
					GENERAL INFORMA
					Sequence 9, Appl
					Sequence 9, Appl
					Sequence 9, Appl
					Sequence 3, Appl
					Sequence 2, Appl
					Sequence 5, Appl
					Sequence 2, Appl
					Sequence 3, Appl
					Sequence 1, Appl
					Sequence 5, Appl
					Sequence 14, Appl
					Sequence 8, Appl
					Sequence 7, Appl
					Sequence 7, Appl
					Sequence 1, Appl

32 645.6 43.0 1303 36 US-09-994-427A-17 Sequence 17, Appl  
33 645.6 43.0 1303 36 US-09-995-419A-6 Sequence 6, Appl1  
34 642.4 42.8 1617 22 US-09-593-316-3 Sequence 3, Appl1  
35 642.4 42.8 1828 7 US-08-188-607B-8 Sequence 8, Appl1  
36 642.4 42.8 1828 5 US-08-378-617-8 Sequence 8, Appl1  
37 642.4 42.8 1828 13 US-08-984-900-8 Sequence 8, Appl1  
38 631.8 42.1 1064 16 US-09-230-091-5 Sequence 5, Appl1  
39 616.2 41.1 1028 16 US-09-230-091-7 Sequence 7, Appl1  
40 583.2 38.9 1353 6 US-08-214-580-1 Sequence 1, Appl1  
41 583.2 38.9 1353 15 US-09-137-517-1 Sequence 1, Appl1  
42 561.8 37.5 1885 22 US-09-593-316-11 Sequence 11, Appl  
43 476.4 31.8 631 17 US-09-371-508-2528 Sequence 2528, Ap  
44 476.4 31.8 631 17 US-09-371-508-2528 Sequence 2528, Ap  
45 476.4 31.8 631 29 US-09-747-508-2528 Sequence 2528, Ap

ALIGNMENTS

RESULT 1  
US-08-379-040-1  
Sequence 1, Application US/08379040  
GENERAL INFORMATION:  
APPLICANT: Cooper, David K.C.  
APPLICANT: Koren, Eugen  
TITLE OF INVENTION: GENETICALLY ENGINEERED ANIMALS FOR USE  
TITLE OF INVENTION: AS ORGAN DONORS  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patrea L. Pabst  
STREET: 1100 Peachtree Street, Suite 2800  
CITY: Atlanta  
STATE: GA  
COUNTRY: USA  
ZIP: 30309-4530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/379,040  
FILING DATE:  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/049,817  
FILING DATE: 20-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: BMC100  
TELEPHONE: (404)-815-6508  
TELEFAX: (404)-815-6555  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1500 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Mus musculus  
PUBLICATION INFORMATION:  
AUTHORS: Larsen, Robert D.  
AUTHORS: Rajan, Valanila P.  
AUTHORS: Ruff, Melissa M.  
AUTHORS: Kukowska-Latallo, Jolanta  
AUTHORS: Cummings, Richard D.  
AUTHORS: Lowe, John B.  
TITLE: Isolation of a cDNA encoding a murine

TITLE: UDP-galactose:beta-D-galactosyl-1,4-N-acetyl-D-  
glucosaminide-alpha-1,3-galactosyltransferase:  
TITLE: Expression cloning by gene transfer.  
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.  
VOLUME: 86  
ISSUE: November  
PAGES: 8227-8231  
DATE: 1989  
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 1500  
US-08-379-040-1

Query Match 100.0%; Score 1500; DB 7; Length 1500;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTTCCCTTTGAGACTCTTCTTGAATGAGAAGTACCGATTCTGCTGAAGACCTCGCGCT 60  
Db 1 CCTTCCCTTTGAGACTCTTCTTGAATGAGAAGTACCGATTCTGCTGAAGACCTCGCGCT 60  
Qy 61 CTCAGGCTCTGGGAGTTGGAAACCTGTACCTTCTCTCTCTGCTGAGCCCTGCTCCTT 120  
Db 61 CTCAGGCTCTGGGAGTTGGAAACCTGTACCTTCTCTCTCTGCTGAGCCCTGCTCCTT 120  
Qy 121 AGGAGGCCACAGCTCGACAGAACTCGGTTGCTTTGCTTTGTTGGAGGGGAACACAG 180  
Db 121 AGGAGGCCACAGCTCGACAGAACTCGGTTGCTTTGCTTTGTTGGAGGGGAACACAG 180  
Qy 181 CTGACGATGAGGCTGACTTTGAACCTCAAGAGATCTGTTACCCAGTCTCTCGGAATTA 240  
Db 181 CTGACGATGAGGCTGACTTTGAACCTCAAGAGATCTGTTACCCAGTCTCTCGGAATTA 240  
Qy 241 AGGCTGTACTACATTTGCCCTGGACCTAAAGATTTTCATGATCACTATGCTTCAAGATCTC 300  
Db 241 AGGCTGTACTACATTTGCCCTGGACCTAAAGATTTTCATGATCACTATGCTTCAAGATCTC 300  
Qy 301 CATGTCAACAGATCTCCATGTCAGATCCAAAGTCCAGATCAGAAACAGTCTTCCATCTCAAGA 360  
Db 301 CATGTCAACAGATCTCCATGTCAGATCCAAAGTCCAGATCAGAAACAGTCTTCCATCTCAAGA 360  
Qy 361 TCTGGATCAGAGGAGAAATTAATCAATGTCAGGAGAAAGTAATCTCTGCTGATGCTGATT 420  
Db 361 TCTGGATCAGAGGAGAAATTAATCAATGTCAGGAGAAAGTAATCTCTGCTGATGCTGATT 420  
Qy 421 GTCTCAACCGTGTGCTGCTTTTGGGAATATGTCACAGAAATTCAGAGGTTGGTGAG 480  
Db 421 GTCTCAACCGTGTGCTGCTTTTGGGAATATGTCACAGAAATTCAGAGGTTGGTGAG 480  
Qy 481 AACAGATGGCAGAGGACTGCTGTTCCCAAGCTGTTTAAATGGGACCCACAGTTAT 540  
Db 481 AACAGATGGCAGAGGACTGCTGTTCCCAAGCTGTTTAAATGGGACCCACAGTTAT 540  
Qy 541 CAAGAAGACACGCTAGAGGACGAGAGAAAGGTAGAAATGGAGATCGCATGAAGAG 600  
Db 541 CAAGAAGACACGCTAGAGGACGAGAGAAAGGTAGAAATGGAGATCGCATGAAGAG 600  
Qy 601 CCTCAGCTATGGGACTGTTTCAATCCAAAGAACCCCGGATGTTTGGACAGTGAACCCG 660  
Db 601 CCTCAGCTATGGGACTGTTTCAATCCAAAGAACCCCGGATGTTTGGACAGTGAACCCG 660  
Qy 661 TGGAGGCGCGATTGCTGGGAAGGACATTTATGACACAGCTCTGCTGGAAAAGTACTAC 720  
Db 661 TGGAGGCGCGATTGCTGGGAAGGACATTTATGACACAGCTCTGCTGGAAAAGTACTAC 720  
Qy 721 GCCACACAGAACTCACTGTGGGCTGACACTGTTTGTGTGGGAAGTACATTCAGCAT 780  
Db 721 GCCACACAGAACTCACTGTGGGCTGACACTGTTTGTGTGGGAAGTACATTCAGCAT 780  
Qy 781 TACTTAGAAGACTTTCTGGAGTCTGCTGACATGCTACTTTCATGGTGGCCATCGGGTCATA 840  
Db 781 TACTTAGAAGACTTTCTGGAGTCTGCTGACATGCTACTTTCATGGTGGCCATCGGGTCATA 840  
Qy 841 TTTTACGTCATGATAGAGGACACCTCCCGGATGCTGCTGTCACCTGAACCCCTCTACAT 900

Db 841 TTTTAGTGCATGATAGACACACCTCCCGGATGCCGTGCTGCACCTGAACCCCTCTACAT 900  
Qy 901 TCCTTACAAGTCTTTGAGATCAGGTCTGAGAAGAGCTGGCAGGATATCAGCATGATGCC 960  
Db 901 TCCTTACAAGTCTTTGAGATCAGGTCTGAGAAGAGCTGGCAGGATATCAGCATGATGCC 960  
Qy 961 ATGAAGACCATTTGGGAGGACATCTCGGCCACATCCAGCAGAGGTCGACTTCTCTTC 1020  
Db 961 ATGAAGACCATTTGGGAGGACATCTCGGCCACATCCAGCAGAGGTCGACTTCTCTTC 1020  
Qy 1021 TCGATGGAGCTGGATCAAGTCTTTCAAGACAATCTCGGGGTGGAATCTCGGCCAGCTG 1080  
Db 1021 TCGATGGAGCTGGATCAAGTCTTTCAAGACAATCTCGGGGTGGAATCTCGGCCAGCTG 1080  
Qy 1081 GTAGCACAGCTCCAGGCTTGGTGTACAAGGCCAGTCCCGAGAGGTTTCACTATGAGAGG 1140  
Db 1081 GTAGCACAGCTCCAGGCTTGGTGTACAAGGCCAGTCCCGAGAGGTTTCACTATGAGAGG 1140  
Qy 1141 CGGGAATCTCGGCCCGCTACATTCCTATTCGGAGAGGGGATTTTACTACACGCGGCC 1200  
Db 1141 CGGGAATCTCGGCCCGCTACATTCCTATTCGGAGAGGGGATTTTACTACACGCGGCC 1200  
Qy 1201 ATTTTGGAGGACGCTACTCAGTCTTCAACCTCAGGAGGAGTCTTAAAGGGATC 1260  
Db 1201 ATTTTGGAGGACGCTACTCAGTCTTCAACCTCAGGAGGAGTCTTAAAGGGATC 1260  
Qy 1261 CTCAGGACAAGAACATCAGATAGAGCCAGTGGCATGTAGAGGCCACCTCAACAAA 1320  
Db 1261 CTCAGGACAAGAACATCAGATAGAGCCAGTGGCATGTAGAGGCCACCTCAACAAA 1320  
Qy 1321 TACTTCTTTTCAACAAACCCACTAAATCTATCTCCAGAGTATTTGCTGGGACTATCAG 1380  
Db 1321 TACTTCTTTTCAACAAACCCACTAAATCTATCTCCAGAGTATTTGCTGGGACTATCAG 1380  
Qy 1381 ATAGGCTGCTTCAGATATTAAGTGTCAAGGTAGCTTGGCAGACAAAGAGTATAAT 1440  
Db 1381 ATAGGCTGCTTCAGATATTAAGTGTCAAGGTAGCTTGGCAGACAAAGAGTATAAT 1440  
Qy 1441 TTGGTTAGAAATTAATCTCGACTTCAAAATTTGATGGAACCTTGACACTATTTCTAACCA 1500  
Db 1441 TTGGTTAGAAATTAATCTCGACTTCAAAATTTGATGGAACCTTGACACTATTTCTAACCA 1500

RESULT 2  
US-08-823-489-3  
; Sequence 3, Application US/08823489  
; GENERAL INFORMATION:  
; APPLICANT: LOWE, JOHN B.  
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUBSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/823,489  
; FILING DATE: 25-MAR-1997  
; CLASSIFICATION: 435  
; PRIORITY DATA:  
; PRIORITY NUMBER: 07/914,281

; FILING DATE: 20-JUL-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lavalleye, Jean-Paul M. P.  
; REGISTRATION NUMBER: 31,451  
; REFERENCE/DOCKET NUMBER: 2363-060-55  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)521-4500  
; TELEFAX: (703)486-2347  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1500 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
; ANTI-SENSE: NO  
; US-08-823-489-3

Query Match 100.0%; Score 1500; DB 12; Length 1500;  
Best Local Similarity 100.0%; Pred No. 0;  
Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTTCCCTTTGATAGACTCTTCTTGGAAATGAGAAGTACCGATTCTGCTGAAGACCTCGCGCT 60  
Db 1 CCTTCCCTTTGATAGACTCTTCTTGGAAATGAGAAGTACCGATTCTGCTGAAGACCTCGCGCT 60  
Qy 61 CTCAGGCTCTGGAGTGGAAACCTGTACCTTCTCTTCTCTCTGCTGAGCCCTGCTTCTT 120  
Db 61 CTCAGGCTCTGGAGTGGAAACCTGTACCTTCTCTTCTCTCTGCTGAGCCCTGCTTCTT 120  
Qy 121 AGCAGGCCAGAGCTCGACAGAACTCGGTGCTTTCGCTTTCGCTTGGAGGGAACACAG 180  
Db 121 AGCAGGCCAGAGCTCGACAGAACTCGGTGCTTTCGCTTTCGCTTGGAGGGAACACAG 180  
Qy 181 CTGACGATGAGGCTGACTTTTGAAGTCAAGAGATCTGCTTACCCAGTCTCCTTGAATATA 240  
Db 181 CTGACGATGAGGCTGACTTTTGAAGTCAAGAGATCTGCTTACCCAGTCTCCTTGAATATA 240  
Qy 241 AGCCCTGTACTACATTTGCTTGGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCTC 300  
Db 241 AGCCCTGTACTACATTTGCTTGGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCTC 300  
Qy 301 CATGTCAACAAGATCTCCATGTCAAGATCCAAGTCAAGAAACAAGTCTTCCATCTCTCAAGA 360  
Db 301 CATGTCAACAAGATCTCCATGTCAAGATCCAAGTCAAGAAACAAGTCTTCCATCTCTCAAGA 360  
Qy 361 TCTGGATCACAGGAGAAAATAATGAATGTCAAGGGAAGAAATAATCTCTGTTGATGCTGATT 420  
Db 361 TCTGGATCACAGGAGAAAATAATGAATGTCAAGGGAAGAAATAATCTCTGTTGATGCTGATT 420  
Qy 421 GTCTCAACCTGCTTCTGCTGTTTGGGAATATGTCAACAGAAATTCAGAGGTTGGTGAG 480  
Db 421 GTCTCAACCTGCTTCTGCTGTTTGGGAATATGTCAACAGAAATTCAGAGGTTGGTGAG 480  
Qy 481 AACAGATGGCAGAGGAGTGGTGTCCCAAGCTGGTTTAAATAATGGAGCCACAGTTAT 540  
Db 481 AACAGATGGCAGAGGAGTGGTGTCCCAAGCTGGTTTAAATAATGGAGCCACAGTTAT 540  
Qy 541 CAAGAAGAACCGTAGAAGGACGAGAGAAAGGGTAGAAATGGAGATCCGATTTGAAGAG 600  
Db 541 CAAGAAGAACCGTAGAAGGACGAGAGAAAGGGTAGAAATGGAGATCCGATTTGAAGAG 600  
Qy 601 CCTCAGCTATGGGACTGGTTTCAATCCAAAGACCGCCCGGATGTTTTCAGAGTGCACCCG 660  
Db 601 CCTCAGCTATGGGACTGGTTTCAATCCAAAGACCGCCCGGATGTTTTCAGAGTGCACCCG 660  
Qy 661 TGGAGGCGCGGATTTGTTGGGAGGACCTTATGACACAGCTCTGCTGGGAAAAGTACTAC 720  
Db 661 TGGAGGCGCGGATTTGTTGGGAGGACCTTATGACACAGCTCTGCTGGGAAAAGTACTAC 720  
Qy 721 GCCACACAGAAACTCAGTGTGGGGCTGACAGTGTGTTGCTGCTGGGAAAAGTACTATGAGCAT 780

Db 721 GCCACAGAAACTCACTGTGGGCTGACAGTGTCTGTGGAAAGTACATTGACCAT 780  
QY 781 TACTTAGAAGACTTCTGGAGTCTGCTGACATGTACTTTCATGTTGGCCATCGGGTCATA 840  
Db 781 TACTTAGAAGACTTCTGGAGTCTGCTGACATGTACTTTCATGTTGGCCATCGGGTCATA 840  
QY 841 TTTTACGTATGATAGACGACACTCCCGGATCCCTGCTGCTGACACTGAACCCCTCTACAT 900  
Db 841 TTTTACGTATGATAGACGACACTCCCGGATCCCTGCTGCTGACACTGAACCCCTCTACAT 900  
QY 901 TCCTTACAAGTCTTTGAGATCAGTCTGAGAAGAGTGGGAGATATCAGCATGATCGGC 960  
Db 901 TCCTTACAAGTCTTTGAGATCAGTCTGAGAAGAGTGGGAGATATCAGCATGATCGGC 960  
QY 961 ATGAAGACCAATTTGGGAGACATCTTGGCCACATCCAGCACGAGGTGCATCTTCTCTTC 1020  
Db 961 ATGAAGACCAATTTGGGAGACATCTTGGCCACATCCAGCACGAGGTGCATCTTCTCTTC 1020  
QY 1021 TGATGACGTGGATCAAGTCTTTTCAAGACAACCTTGGGGTGGAAACTCTGGGCCAGCTG 1080  
Db 1021 TGATGACGTGGATCAAGTCTTTTCAAGACAACCTTGGGGTGGAAACTCTGGGCCAGCTG 1080  
QY 1081 GTAGCAGCTCCAGGCTCGTGTGTACAAAGCCAGTCCCGAGAGTTTCACTATGAGAGG 1140  
Db 1081 GTAGCAGCTCCAGGCTCGTGTGTACAAAGCCAGTCCCGAGAGTTTCACTATGAGAGG 1140  
QY 1141 CGGGAATCTCGGCGCGGTACATTTCCATTCGAGAGGGGATTTTACTACCAACCGCGCC 1200  
Db 1141 CGGGAATCTCGGCGCGGTACATTTCCATTCGAGAGGGGATTTTACTACCAACCGCGCC 1200  
QY 1201 ATTTTGGAGAACGCTACTACATCTCAACCTCACAGGAGTCTTTAAAGGGATC 1260  
Db 1201 ATTTTGGAGAACGCTACTACATCTCAACCTCACAGGAGTCTTTAAAGGGATC 1260  
QY 1261 CTCAGGACAAGAACATGATAGAGCCAGTGGCATGATGAGGCCACCTCAACAAA 1320  
Db 1261 CTCAGGACAAGAACATGATAGAGCCAGTGGCATGATGAGGCCACCTCAACAAA 1320  
QY 1321 TACTTCTCTTTTCAACAAACCCACTAAAATCCTATCTCCAGAGTATTTGCTGGGACTATCAG 1380  
Db 1321 TACTTCTCTTTTCAACAAACCCACTAAAATCCTATCTCCAGAGTATTTGCTGGGACTATCAG 1380  
QY 1381 ATAGGCTGCTTTCAGATATTTAAAGTGTCAAGGTAGCTTGGCAGACAAAAGATATAAT 1440  
Db 1381 ATAGGCTGCTTTCAGATATTTAAAGTGTCAAGGTAGCTTGGCAGACAAAAGATATAAT 1440  
QY 1441 TTGGTTAGAAATATGCTGACTTCAAAATTTGATGGAACCTTGACACTATTTCTAACCA 1500  
Db 1441 TTGGTTAGAAATATGCTGACTTCAAAATTTGATGGAACCTTGACACTATTTCTAACCA 1500

RESULT 3

US-09-593-316-9  
; Sequence 9, Application US/09593316  
; GENERAL INFORMATION:  
; APPLICANT: Clark Mr., John  
; APPLICANT: Denning, Chris  
; TITLE OF INVENTION: Animal Tissue For Xenotransplantation  
; FILE REFERENCE: 730/002  
; CURRENT APPLICATION NUMBER: US/09/593,316  
; PRIOR FILING DATE: 2000-06-13  
; PRIOR FILING DATE: 2000-05-15  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 1500  
; TYPE: DNA  
; ORGANISM: Musca sp.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (277)..(1461)

US-09-593-316-9

Query Match 100.0%; Score 1500; DB 22; Length 1500;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCTTCCCTTGTAGACTCTTCTTGGAAATGAGAATGACGATTTCTGCTCAAGACCTCGCGCT 60  
Db 1 ccttccctctgtagactctctcttggaatgagaagtaacgattctgctgaagacctcgagct 60  
QY 61 CTCAGGCTCTGGGAGTGTGGAAACCTGTACCTTCTCTTTCCTCTGCTGAGCCCTGCTCTT 120  
Db 61 ctcaggctctgggagttggaaacctgtaccttctctctctctctctgctgagccctgctcctt 120  
QY 121 AGGAGCCGACAGAGCTGCACAGAACTCGGTTGCTTGTGTTTGTGTTTGGAGGGAACACAG 180  
Db 121 aggagccgagagctgcacagaaactcgtctgctctgctctgctcttgaggggaacacag 180  
QY 181 CTGACGATGAGGCTGACTTTTGAACCTCAAGAGATCTGCTTACCCAGTCTCTTGGAAATTA 240  
Db 181 ctgacgatgaggtctgactttgaaactcaagagatctgcttaccctcagctctctctggaattaa 240  
QY 241 AGGCTGTACTACATTTGCTTGGACCTTAAGATTTTCATGATCAGTATGCTTCAAGATCTC 300  
Db 241 aggcctgtactacatttgcctggacctaaagtcttcatgatacactatgcttcaagatctc 300  
QY 301 CATGTCAACAAAGATCTCCATGTCAAGATCCAAGTCAAGAAACAGTCTTCCATCTCTCAAGA 360  
Db 301 catgtcaacaagaatctccatgtccaagtccaagtcagaacaagaatcttccatctctcaaga 360  
QY 361 TCTGGATCACAGGAGAAATAATGAATGTCAAGGGAAGAAAGTAATCTCTTGTGATGCTGAT 420  
Db 361 tctgga tcacagggagaaaaataatgaatgtccaagggaaaagtaacctctgtgtgctgatt 420  
QY 421 GTCTCAACCGTGTGTGCTGCTTGTGGGAATATGTCAACAGAAATTCAGAGGTTGGTGAG 480  
Db 421 gtctcaaccgtgtgtgctgcttgtgggaatattgtccaagaaatccacagaaatccacagaggtggtag 480  
QY 481 AACAGATGGGAGAGGACTGTGTGCTTCCCAAGCTGGTTTTAAATAATGGGACCCACAGTTAT 540  
Db 481 aacaga tggcagaagaactggtgtgtcccaagctggtttcaaaaaatgggacccacagttat 540  
QY 541 CAAGAAGACAACGTAGAGGACGAGAGAAAGGTTAGAAATGGAGATCGGATTTGAAGAG 600  
Db 541 caagaagacaacgtagaaggacgagagaaaggttagaaatggagatcgcatttgaagag 600  
QY 601 CCTCAGCTATGGGACTGTTCAATCCAAGAACCCCGGATGTTTTCACAGTGCACCCG 660  
Db 601 cctcagcta tgggactggttcaa tcccaagaaccccgagatgttttgacagtgaccctcg 660  
QY 661 TGAAGGCGCGGATTTGTGTGGAAAGGCACTTATGACACAGCTCTGCTGGAAGAGTACTAC 720  
Db 661 tggaaagcgccgattgtgtgggaaggcacttatgacacagctctgctggaaaagtactac 720  
QY 721 GCCACAGAAACTCACTGTGGGCTGACAGTGTGTTGCTGTGGGAAAGTACATTTGAGCAT 780  
Db 721 gccacagaaactcactgtgggctgacagttgtgtgtgggaaagtaacattgagcat 780  
QY 781 TACTTAGAAGACTTCTTGGAGTCTGCTGACATGTACTTTCATGTTGGCCATCGGGTCATA 840  
Db 781 tacttagaagactctctggagctctgctgacatgtacttca tgg ttggcca tggggtcata 840  
QY 841 TTTTACGTATGATAGACGACACTCCCGGATCCCTGCTGCTGACACTGAACCCCTCTACAT 900  
Db 841 ttttacgtatgatagacgacacctcccgatgctgtctgacacctgaacctctacat 900  
QY 901 TCCTTACAAGTCTTTGAGATCAGTCTGAGAAGAGTGGGAGATATCAGCATGATCGGC 960  
Db 901 tccttaaaagctcttggatgatacaggtctgagaagaggtgggagagatacagcatgatgcgc 960  
QY 961 ATGAAGACCAATTTGGGAGGACACTCTGCGCCACATCTCAGACACGAGGTGCTCTCTTC 1020  
Db 961 atgaagaccaatttgggagggacactctgcccacatctcagacacgaggtgctctctcttc 1020

Db 961 atgaagaccattgaggagacacatccttgccccacatccacagcagaggtcgacttctcttc 1020  
Qy 1021 TCATCGAGCGTGGATCAAGTCCTTCAAGACAACTTCGGGGTGGAACTCTGGCCGAGCTG 1080  
Db 1021 tgcagagctgagatcaagctcttcaagaacaacttcggggtggaactctggccagctg 1080  
Qy 1081 GTAGCACAGCTCCAGCGCTGGGTGTACAGGCCAGTCCCGAGAAAGTTACCTATGAGAG 1140  
Db 1081 gLagcacagctccaggcctggTggLacaaggccagTcccgaagaagTccacctatgagag 1140  
Qy 1141 CGGGAACCTCGCGCGCTGACATTCATTCGAGAGGGGGATTTTACTACCACCGCGCC 1200  
Db 1141 cgggaacTgcgcgcgcgcacattccattcgcgagaggggatttttactaccacgcgcgc 1200  
Qy 1201 ATTTTGGAGGACCGCTACTCACATTTCTAACCTCACCGGAGTGCCTTAAGGGGATC 1260  
Db 1201 attttggaggaacgcctactcacattctcaacctcaccagagtgcttgaaggatc 1260  
Qy 1261 CTCAGGACAAGAATGACATAGAGCCCGAGTGCGATGATGAGAGCCACCTCAACAAA 1320  
Db 1261 ctccaggacaagaacatgacatagaagcccgatggcatgatgagagccacctcaacaaa 1320  
Qy 1321 TACTTCCCTTTCAACAAACCCACTAAATCCTATCTCCAGACTATTGCTGGGACTATCAG 1380  
Db 1321 tacttctctttcaacaacccccactaaaatctctatctccagagtattgctgggactatcg 1380  
Qy 1381 ATAGGCTGCCCTTCAGATATTAAGTGTCAAGGTAGCTTGGCAGACAAAAGAGTATAAT 1440  
Db 1381 atagcctgcctcagatattaaaagtgcaagtgatgctggcagaaaaagagataat 1440  
Qy 1441 TTGGTTAGAAATAATGCTGACTTCAAAATGTGATGGAACCTTGACACTATTTCCTAAACA 1500  
Db 1441 ttggttagaaaTaatgtcgtacttcaaatgtgataggaaacttgacactatttctlaacca 1500

RESULT 4  
US-09-863-475-3  
; Sequence 3, Application US/09863475  
; GENERAL INFORMATION:  
; APPLICANT: LOWE, JOHN B.  
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
; GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
; OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/863,475  
; FILING DATE: 24-May-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/914,281  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lavalleye, Jean-Paul M. P.  
; REGISTRATION NUMBER: 31,451  
; REFERENCE/DOCKET NUMBER: 2363-060-55  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)521-4500  
; TELEFAX: (703)486-2347  
; TELEX: 248855 OPAT UR

; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1500 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
; ANTI-SENSE: NO  
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-863-475-3

Query Match 100.0%; Score 1500; DB 33; Length 1500;  
Best Local Similarity 100.0%; Pred. NO. 0;  
Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

Qy 1 CCTTCCCTTGTAGACTCTTCTTGGAAATGAGAAGTACCGATTCTGCTGAAGACCTCGCGCT 60  
Db 1 CCTTCCCTTGTAGACTCTTCTTGGAAATGAGAAGTACCGATTCTGCTGAAGACCTCGCGCT 60  
Qy 61 CTCAGGCTCTGGGAGTGTGAACCCCTGTAGCTTCTCTTTCCTCTCTGCTGAGCCCTGCCCTCTT 120  
Db 61 CTCAGGCTCTGGGAGTGTGAACCCCTGTAGCTTCTCTTTCCTCTCTGCTGAGCCCTGCCCTCTT 120  
Qy 121 AGCAGGCCAGAGCTCGACAGAACTCGGTTCGCTTTCCTCTGCTGTTGGAGGAAACACAG 180  
Db 121 AGCAGGCCAGAGCTCGACAGAACTCGGTTCGCTTTCCTCTGCTGTTGGAGGAAACACAG 180  
Qy 181 CTGACGATGAGCTGACTTGAACCTCAAGAGATCTGCTTACCCAGTCTCTCTGGAATTA 240  
Db 181 CTGACGATGAGCTGACTTGAACCTCAAGAGATCTGCTTACCCAGTCTCTCTGGAATTA 240  
Qy 241 AGSCCTGTACTACATTTGCTTGGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCTC 300  
Db 241 AGSCCTGTACTACATTTGCTTGGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCTC 300  
Qy 301 CATGTCAACAAGATCTCCATGCAAGATCCAACTCAGAACTCAAGATCTTCATCTCTCAAGA 360  
Db 301 CATGTCAACAAGATCTCCATGCAAGATCCAACTCAGAACTCAAGATCTTCATCTCTCAAGA 360  
Qy 361 TCTGGATCAGAGAGAAATTAATGAATGTCAGGGAAAGTAACTCTGTTGATGCTGATTT 420  
Db 361 TCTGGATCAGAGAGAAATTAATGAATGTCAGGGAAAGTAACTCTGTTGATGCTGATTT 420  
Qy 421 GTCTCAACCTGCTGCTGCTGTTTGGGAATATGTCAACAGAATTCACAGAGTTGGTGAG 480  
Db 421 GTCTCAACCTGCTGCTGCTGTTTGGGAATATGTCAACAGAATTCACAGAGTTGGTGAG 480  
Qy 481 AACAGATGGCAGAAAGACTGGTGGTTCCTCAAGCTGGTTTAAAAATGGGACCCACAGTTAT 540  
Db 481 AACAGATGGCAGAAAGACTGGTGGTTCCTCAAGCTGGTTTAAAAATGGGACCCACAGTTAT 540  
Qy 541 CAAGAAGACAACTAGAAAGCGGAGAGAAAGGGTAGAANTGGAGTCCGATTTGAAGAG 600  
Db 541 CAAGAAGACAACTAGAAAGCGGAGAGAAAGGGTAGAANTGGAGTCCGATTTGAAGAG 600  
Qy 601 CCTCAGCTATGGGACTGGTTCAATCCAAAGAACCCCGCGGATGTTTTGACAGTGACCCCG 660  
Db 601 CCTCAGCTATGGGACTGGTTCAATCCAAAGAACCCCGCGGATGTTTTGACAGTGACCCCG 660  
Qy 661 TGAAGGCCCGGATTTGTTGGGAAAGGACTTATGACACAGCTCTGCTGGGAAAGTACTAC 720  
Db 661 TGAAGGCCCGGATTTGTTGGGAAAGGACTTATGACACAGCTCTGCTGGGAAAGTACTAC 720  
Qy 721 GCCACACAGAACTCACTGTGGGGCTGACAGTCTTTCCTGCTGGGAAAGTACTACAGCAT 780  
Db 721 GCCACACAGAACTCACTGTGGGGCTGACAGTCTTTCCTGCTGGGAAAGTACTACAGCAT 780  
Qy 781 TACTTGAAGACTTTCTGAGTCTGCTGACATGACTTACTTTCCTGCTGGGAAAGTACTACAT 840  
Db 781 TACTTGAAGACTTTCTGAGTCTGCTGACATGACTTACTTTCCTGCTGGGAAAGTACTACAT 840  
Qy 841 TTTTACGTGATGATAGACGACACCTCCCGGATGCTGCTGCACTGAAACCTCTTACAT 900





QY 721 GCCACACAGAACTACTGTGGGCTGACAGTGTTCGTGTGGAAAGTACATTGAGCAT 780  
Db 721 GCCACACAGAACTACTGTGGGCTGACAGTGTTCGTGTGGAAAGTACATTGAGCAT 780  
QY 781 TACTTAGAGACATTTCTGGAGTCTGTGACATCTACTTCTATGTTGGCCATCGGTCATA 840  
Db 781 TACTTAGAGACATTTCTGGAGTCTGTGACATCTACTTCTATGTTGGCCATCGGTCATA 840  
QY 841 TTTTACGTCATGATAGACACACCTCCGGATGCCTGTCTGTCACCTGAAACCTCTACAT 900  
Db 841 TTTTACGTCATGATAGACACACCTCCGGATGCCTGTCTGTCACCTGAAACCTCTACAT 900  
QY 901 TCCTTACAAGCTTTGAGATCAGGTCTGAGAAGAGTGGCAGGATATCAGCATGATGCCG 960  
Db 901 TCCTTACAAGCTTTGAGATCAGGTCTGAGAAGAGTGGCAGGATATCAGCATGATGCCG 960  
QY 961 ATCAAGACCAATGGGAGCAGCATCTGGCCACATCCAGCAGGTCGACTTCTCTTC 1020  
Db 961 ATCAAGACCAATGGGAGCAGCATCTGGCCACATCCAGCAGGTCGACTTCTCTTC 1020  
QY 1021 TGCATGACGTGATCAAGTCTTTCAAGACAACCTTCGGGTGGAACCTCTGGGCCAGCTG 1080  
Db 1021 TGCATGACGTGATCAAGTCTTTCAAGACAACCTTCGGGTGGAACCTCTGGGCCAGCTG 1080  
QY 1081 GTAGCAGCTCAGGCTGTGTGTACAGGCCAGTCCCGAGAGTTCACCTATGAGAGG 1140  
Db 1081 GTAGCAGCTCAGGCTGTGTGTACAGGCCAGTCCCGAGAGTTCACCTATGAGAGG 1140  
QY 1141 CGGGAAGTCTCGGCCGCTACATCTCCATTCGGAGAGGGGATTTTACTACACGCGGCC 1200  
Db 1141 CGGGAAGTCTCGGCCGCTACATCTCCATTCGGAGAGGGGATTTTACTACACGCGGCC 1200  
QY 1201 ATTTTGGAGAACGCTACTACATCTCAACCTTCACAGGAGTGTCTTAAGGGGATC 1260  
Db 1201 ATTTTGGAGAACGCTACTACATCTCAACCTTCACAGGAGTGTCTTAAGGGGATC 1260  
QY 1261 CTCAGGACAAGAACATGACATAGAACGCCAGTGGGATGATGAGAGCCACCTCAACAAA 1320  
Db 1261 CTCAGGACAAGAACATGACATAGAACGCCAGTGGGATGATGAGAGCCACCTCAACAAA 1320  
QY 1321 TACTTCTTTCAACAAACCCACTAAATCTATCTCCAGAGTATTCCTGGGACTATCAG 1380  
Db 1321 TACTTCTTTCAACAAACCCACTAAATCTATCTCCAGAGTATTCCTGGGACTATCAG 1380  
QY 1381 ATAGGCTGCTTTCAGATATTAAGTGTCAAGTGTCAAGTGTGGCAGACAAAGAGTATAAT 1440  
Db 1381 ATAGGCTGCTTTCAGATATTAAGTGTCAAGTGTCAAGTGTGGCAGACAAAGAGTATAAT 1440  
QY 1441 TTGGTTAGAAATATGTCTGACTTCAAAATGTGATGGAACCTTGACACTATTTCTAACCA 1500  
Db 1441 TTGGTTAGAAATATGTCTGACTTCAAAATGTGATGGAACCTTGACACTATTTCTAACCA 1500

RESULT 6

US-09-946-034-1

GENERAL INFORMATION:

APPLICANT: Cooper, David K.C.  
Koren, Eugen

TITLE OF INVENTION: GENETICALLY ENGINEERED ANIMALS FOR USE

AS ORGAN DONORS

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Patrea L. Pabst

STREET: 1100 Peachtree Street, Suite 2800

CITY: Atlanta

STATE: GA

COUNTRY: USA

ZIP: 30309-4530

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/946,034  
FILING DATE: 04-Sep-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/049,817  
FILING DATE: 20-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: BMC100  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404)-815-6508  
TELEFAX: (404)-815-6555  
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.  
VOLUME: 86  
ISSUE: November  
PAGES: 8227-8231  
DATE: 1989  
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 1500  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-946-034-1

Query Match 100.0%; Score 1500; DB 35; Length 1500;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTCCCTTCTAGACTCTCTTGGAAATGAGAACTGACGATTCCTGCTGAGCCCTGCTCTCTT 60  
Db 1 CTTTCCCTTCTAGACTCTCTTGGAAATGAGAACTGACGATTCCTGCTGAGCCCTGCTCTCTT 60  
QY 61 CTCAGGCTCTGGGAGTGTGGAAACCTGTACCTTCTCTCTCTGCTGAGCCCTGCTCTCTT 120  
Db 61 CTCAGGCTCTGGGAGTGTGGAAACCTGTACCTTCTCTCTCTGCTGAGCCCTGCTCTCTT 120  
QY 121 AGGAGGCCAGAGCTCGACAGAACTCGGTTGCTGTTGCTGTTGCTGTTGAGGGAGAACAG 180  
Db 121 AGGAGGCCAGAGCTCGACAGAACTCGGTTGCTGTTGCTGTTGCTGTTGAGGGAGAACAG 180  
QY 181 CTGACGATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCAGTCTCCTCGAAATTA 240  
Db 181 CTGACGATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCAGTCTCCTCGAAATTA 240  
QY 241 AGGCTGTACTTACATTTGCTGACCTTAAGATTTTTCATGATCAGTCTGCTTCAAGATCTC 300  
Db 241 AGGCTGTACTTACATTTGCTGACCTTAAGATTTTTCATGATCAGTCTGCTTCAAGATCTC 300  
QY 301 CATGTCAACAGATCTCCATGTCAAGATCCAAAGTCAAGTCAAGAAACAGTCTTCCATCTCA 360  
Db 301 CATGTCAACAGATCTCCATGTCAAGATCCAAAGTCAAGTCAAGAAACAGTCTTCCATCTCA 360  
QY 361 TCTGGATCAGAGGAGAAATTAATGAATGTCAAGGAAAGTAATCTCTGTTGATGCTGATT 420  
Db 361 TCTGGATCAGAGGAGAAATTAATGAATGTCAAGGAAAGTAATCTCTGTTGATGCTGATT 420  
QY 421 GTCTCAACCGTGTGCTGCTGTTTGGGAATATGTCAACAGAAATTCAGAGGTTGGTCAG 480  
Db 421 GTCTCAACCGTGTGCTGCTGTTTGGGAATATGTCAACAGAAATTCAGAGGTTGGTCAG 480  
QY 481 AACAGATGGCAGAGGAGTGGTGGTCCCAAGCTGGTTTAAATGGGACCCACAGTTAT 540  
Db 481 AACAGATGGCAGAGGAGTGGTGGTCCCAAGCTGGTTTAAATGGGACCCACAGTTAT 540  
QY 541 CAAGAAGACACCTGAGAGGAGGAGAAAGGTTAGAAATGAGATCGCATTTGAAGAG 600  
Db 541 CAAGAAGACACCTGAGAGGAGGAGAAAGGTTAGAAATGAGATCGCATTTGAAGAG 600  
QY 601 CCTCAGCTATGGGACTGTTCAATCCAAAGAACCCCGCGGATGTTTTCACAGTCAACCCCG 660  
Db 601 CCTCAGCTATGGGACTGTTCAATCCAAAGAACCCCGCGGATGTTTTCACAGTCAACCCCG 660  
QY 661 TGGAGGCCCGATTTGTGTGGGAAGGCACTTATGACACAGCTCTGCTGGAAAGTACTTAC 720



INFORMATION FOR SEQ ID NO: 9:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3450  
 TYPE: nucleic acid  
 STRANDEDNESS: double stranded  
 TOPOLOGY: linear  
 MOLECULE TYPE: CDNA  
 DESCRIPTION:  
 HYPOTHETICAL: no  
 ANTI-SENSE: no  
 FRAGMENT TYPE:  
 ORIGINAL SOURCE:  
 ORGANISM: Mus musculus  
 STRAIN:  
 INDIVIDUAL ISOLATE:  
 IMMEDIATE SOURCE:  
 LIBRARY:  
 CLONE:  
 POSITION IN GENOME:  
 CHROMOSOME/SEGMENT:  
 FEATURE:  
 NAME/KEY: Nucleotide sequence for mouse alpha-1-3-galactosyltransferase g  
 PUBLICATION INFORMATION:  
 AUTHORS: Joziassse, D. H.  
 AUTHORS: Shaper, N.L.  
 AUTHORS: Kim, D.  
 AUTHORS: Van den Eljnden, D. H.  
 AUTHORS: Shaper, J. H.  
 TITLE: Murine alpha-1,3-galactosyltransferase: A single gene locus specif  
 JOURNAL: The Journal of Biological Chemistry  
 VOLUME: 267  
 ISSUE: 8  
 PAGES: 5534-5541  
 DATE: 1992  
 US-08-188-607B-g

0 - 1

Query Match		93.9%	Score 1408.4	DB 5	Length 3450
Best Local Similarity		96.8%	Prod. No. 0		
Matches 1483		Conservative	0	Mismatches 11	Indels 38
				Gaps	3

  

Qy	1	CCTTCCCTTGAGACTCTTCTGGAA	TGGAAGTACCGATTCTGCTGAAGAC	CTCGCGT	60
Db					
	64	CCTTCCCTTGAGACTCTTCTGGAA	TGGAAGTACCGATTCTGCTGAAGAC	CTCGCGT	123
Qy	61	CTCAGGCTCTGGGAGTTGGAACCC	TT-GTACCTTCCCTTCTCTGCTGAGCCCTGCCTCT	119	
Db					
	124	CTCAGGCTCTGGGAGTTGGAACCC	TCGTACCTTCTCTGCTGAGCCCTGCCTCT	183	
Qy	120	TAGGAGGCCAGAGCTCGACAGAA	CTCGTTCGTTGCTTGTGCTTGGAGGGAACACA	179	
Db					
	184	TCGGAGGCCAGAGCTGACAGAA	CTCGTTCGTTGCTTGTGCTTGGAGGGAACACA	243	
Qy	180	GCTGACGATGAGGCTGACTTTGAA	CTCAAGACATCTGCTTACCCCAAGTCTCTTGGAA	239	
Db					
	244	GCTGACGATGAGGCTGACTTTGAA	CTCAAGACATCTGCTTACCCCAAGTCTCTTGGAA	303	
Qy	240	AAGGCTCTACTACATTTTGCCT	TGGACCTTAAGATTTTTCATGATCACTATGCTTCAAGATCT	299	
Db					
	304	AAGGCTCTACTAC-CTTTG	CCCTGGACCTTAAGATTTTTCATGATCACTATGCTTCAAGATCT	362	
Qy	300	CCATGCTCAACAAGATCTCCAT	GTCCAAGTCCAAGTTCAGAAACAAGTCTTCCATCTCTCAAG	359	
Db					
	363	CCATGCTCAACAAGATCTCCAT	GTCCAAGTCCAAGTTCAGAAACAAGTCTTCCATCTCTCAAG	422	
Qy	360	ATCTGGATCACAGGAAATAAT	GAATGCTCAAGGGAAGTAATCTCTTGAATGCTGAT	419	
Db					
	423	ATCTGGATCACAGGAAATAAT	GAATGCTCAAGGGAAGTAATCTCTTGAATGCTGAT	482	
Qy	420	TGCTCAACCGTGTTGTCGTG	TTTTGGGAATAATGTCAACAG-----	461	
Db					
	483	TGCTCAACCGTGTTGTCGTG	TTTTGGGAATAATGTCAACAGCCCAAGACGGCTCTTCTCTT	542	

```
Qy 462 -----AATTCAGAGTGTGGTGAGAACAGATGGCAGAGGACTGGTG 503
Db 543 GTGGATATATACACAAAAATCCAGAGTGTGGTGAGAACAGATGGCAGAGGACTGGTG 602
Qy 504 GTTCCCAAGCTGGTTTAAAAATGGACCCACAGTTATCAAGAACACACAGCTAGAGGACG 563
Db 603 GTTCCCAAGCTGGTTTAAAAATGGACCCACAGTTATCAAGAACACACAGCTAGAGGACG 662
Qy 564 GAGAGAAAGGTTAGAAATGGAGATCGCATTTGAAGAGCCTCAGCTATGGGACTGGTTCAA 623
Db 663 GAGAGAAAGGTTAGAAATGGAGATCGCATTTGAAGAGCCTCAGCTATGGGACTGGTTCAA 722
Qy 624 TCCAAAGAACGCCCGGATGTTTGGACAGTGACCCCGTGAAGCGCGGATGTTGTGGGA 683
Db 723 TCCAAAGAACGCCCGGATGTTTGGACAGTGACCCCGTGAAGCGCGGATGTTGTGGGA 782
Qy 684 AGGCACCTTATGACACAGCTCTGCTGGAAAGTACTACGCCACACAGAACTCACTGTGGG 743
Db 783 AGGCACCTTATGACACAGCTCTGCTGGAAAGTACTACGCCACACAGAACTCACTGTGGG 842
Qy 744 GCTGACAGTGTGCTGTGGGAAAGTACATTTAGCATTTACTTTAGAAGACTTTCTGGAGTC 803
Db 843 GCTGACAGTGTGCTGTGGGAAAGTACATTTAGCATTTACTTTAGAAGACTTTCTGGAGTC 902
Qy 804 TGCTGACATGTACTTTCATGTTGGCCATCGGGTGCATATTTTACGTCATGATAGACGAC 863
Db 903 TGCTGACATGTACTTTCATGTTGGCCATCGGGTGCATATTTTACGTCATGATAGACGAC 962
Qy 864 CTCCCGGATGCTCTGCTGACAGCTGAACCTCTACATTCCTTACAAGTCTTTGAGATCAG 923
Db 963 CTCCCGGATGCTCTGCTGACAGCTGAACCTCTACATTCCTTACAAGTCTTTGAGATCAG 1022
Qy 924 GTCTGAGAAGAGTGGCAGGATATCAGCATGATCGCATGAAGACCATTTGGGGAGCACAT 983
Db 1023 GTCTGAGAAGAGTGGCAGGATATCAGCATGATCGCATGAAGACCATTTGGGGAGCACAT 1082
Qy 984 CTGGGCCCATCAGACAGAGTGCAGTTCCTCTTCTGATGAGACGTGGATCAAGTCTT 1043
Db 1083 CTGGGCCCATCAGACAGAGTGCAGTTCCTCTTCTGATGAGACGTGGATCAAGTCTT 1142
Qy 1044 TCAAGACAACTTCGGGGTGAAGTCTGGGCGAGCTGGTAGCAGAGCTCCAGGCGCTGGTG 1103
Db 1143 TCAAGACAACTTCGGGGTGAAGTCTGGGCGAGCTGGTAGCAGAGCTCCAGGCGCTGGTG 1202
Qy 1104 GTAAAGGCCAGTCCCAGAGATTTACCTATGAGAGCGGGAAGTGTGGCGCGGTACAT 1163
Db 1203 GTAAAGGCCAGTCCCAGAGATTTACCTATGAGAGCGGGAAGTGTGGCGCGGTACAT 1262
Qy 1164 TCCATTCCGAGAGGGGATTTTACTACCACGCGGCCATTTTTCGAGGAAGCCCTACTCA 1223
Db 1263 TCCATTCCGAGAGGGGATTTTACTACCACGCGGCCATTTTTCGAGGAAGCCCTACTCA 1322
Qy 1224 CATTTCAACCTCACAGGAGTCTTTTAAGGGATCTCTCAGGACAAAGAAACATGACAT 1283
Db 1323 CATTTCAACCTCACAGGAGTCTTTTAAGGGATCTCTCAGGACAAAGAAACATGACAT 1382
Qy 1284 AGAAGCCAGTGGCATGATGAGAGCCACCTCAACAAATACTTCTTTTCAACAAACCCAC 1343
Db 1383 AGAAGCCAGTGGCATGATGAGAGCCACCTCAACAAATACTTCTTTTCAACAAACCCAC 1442
Qy 1344 TAAATCCTATCTCCAGATGTTGCTGGGACTATCAGATAGGCGCTGCCTTCAGATATTA 1403
Db 1443 TAAATCCTATCTCCAGATGTTGCTGGGACTATCAGATAGGCGCTGCCTTCAGATATTA 1502
Qy 1404 AAGTGTCAAGGTAGCTTGGCAGACAAAGAGTATAATTTGGTTAAGAAATATGTCGACT 1463
Db 1503 AAGTGTCAAGGTAGCTTGGCAGACAAAGAGTATAATTTGGTTAAGAAATATGTCGACT 1562
Qy 1464 TCAATTTGTGATGGAAACTTTGACACTATTTCT 1495
Db 1563 TCAATTTGTGATGGAAACTTTGACACTATTTCT 1594
```

```
RESULT 9
US-08-378-617-9
: Sequence 9, Application US/08378617
: GENERAL INFORMATION:
: APPLICANT: d'Apice, Anthony J.F.
: APPLICANT: Pearse, Martin J.
: APPLICANT: Robins, Allan J.
: APPLICANT: Crawford, Robert J.
: APPLICANT: Rathjen, Peter D.
: TITLE OF INVENTION: MATERIALS AND METHODS FOR MANAGEMENT OF
: TITLE OF INVENTION: HYPERACUTE REJECTION IN HUMAN XENOTRANSPLANTATION
: NUMBER OF SEQUENCES: 33
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson
: STREET: 120 South Sixth Street, Suite 2500
: CITY: Minneapolis
: STATE: MN
: COUNTRY: USA
: ZIP: 55402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30B
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/378,617
: FILING DATE: 26-JAN-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Ellinger, Mark S.
: REGISTRATION NUMBER: 34,812
: REFERENCE/DOCKET NUMBER: 06868/005001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (612) 335-5070
: TELEFAX: (612) 288-9696
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3450 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA
US-08-378-617-9
```

```
Query Match 93.9% Score 1408.4; DB 7; Length 3450;
Best Local Similarity 96.8% Pred. No. 0;
Matches 1483; Conservative 0; Mismatches 11; Indels 38; Gaps 3;
```

```
Qy 1 CCTTCCCTTGTAGACTCTTCTTGGAAATGAGAAGTACCGATTCTGCTGAAGACCTCGCGCT 60
Db 64 CCTTCCCTTGTAGACTCTTCTTGGAAATGAGAAGTACCGATTCTGCTGAAGACCTCGCGCT 123
Qy 61 CTCAGGCTCTGGAGTGTGAACCCCT-GTACCTTCTCTTCTCTGCTGAGGCGCTGCTCTCT 119
Db 124 CTCAGGCTCTGGAGTGTGAACCCCTCGTACCTTCTCTTCTGCTGCTGAGGCGCTGCTCT 183
Qy 120 TAGCAGGCCAGAGCTCGACAGAACTCGGTTGCTTTTGTCTGTTTGGAGGGAACACA 179
Db 184 TCGGCAGGCCAGAGCTGCACAGAGCTCGGTTGCTTTGCTGTTTGGAGGGAACACA 243
Qy 180 GCTGACGATGAGGCTGACTTTTGAAGTCAAGAGATCTGCTTACCCAGTCTCTTGAATTA 239
Db 244 GCTGACGATGAGGCTGACTTTTGAAGTCAAGAGATCTGCTTACCCAGTCTCTTGAATTA 303
Qy 240 AAGCCCTGCTACTACATTTGCCCTGGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCT 299
Db 304 AAGCCCTGCTACTAC-CCTTGGCTGGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCT 362
Qy 300 CCATGTCAACAAGATCTCCATGTCAAGATCCAAGTCCAAGTCAAGTCAAGTCAAGTCAAG 359
Db 363 CCATGTCAACAAGATCTCCATGTCAAGTCCAAGTCCAAGTCAAGTCAAGTCAAGTCAAG 422
```

QY 360 ATCTGGATCACAGGAGAAATAATGAATGTCAGAGGAAAGTAATCTCTTTGATGCTGAT 419  
DB 423 ATCTGGATCACAGGAGAAATAATGAATGTCAGAGGAAAGTAATCTCTTTGATGCTGAT 482  
QY 420 TCTCTCAACCGTGTGTGCTGTTTGGGAATATGTCACAG----- 461  
DB 483 TGTCTCAACCGTGTGTGCTGTTTGGGAATATGTCACAGCCCAAGACGGCTCTTTCTTT 542  
QY 462 -----AATTCAGAGGTTGTTGAGAACAGATGGCAGAGGACTGGTG 503  
DB 543 GTGGATATATACACAAAAATCCAGAGGTTGTTGAGAACAGATGGCAGAGGACTGGTG 602  
QY 504 GTTCCCAAGCTGTTTAAATAATGGGACCCACAGATTATCAAGAGACAACTAGAGGAGC 563  
DB 603 GTTCCCAAGCTGTTTAAATAATGGGACCCACAGATTATCAAGAGACAACTAGAGGAGC 562  
QY 564 GAGAGAAAAGGTAGAAATGGAGATCGCATTCGAAGAGCCTCAGCTATGGAGCTGGTTCAA 623  
DB 663 GAGAGAAAAGGTAGAAATGGAGATCGCATTCGAAGAGCCTCAGCTATGGAGCTGGTTCAA 722  
QY 624 TCCAAGAACCCCGCGATGTTTTCACAGTGACCCCGTGGAAAGCGCCGATTGTGTGGGA 683  
DB 723 TCCAAGAACCCCGCGATGTTTTCACAGTGACCCCGTGGAAAGCGCCGATTGTGTGGGA 782  
QY 684 AGGCACATTATGACACAGCTCTCTGTGAAAAGTACTACGCCACACAGAAACTCACTGTGG 743  
DB 783 AGGCACATTATGACACAGCTCTCTGTGAAAAGTACTACGCCACACAGAAACTCACTGTGG 842  
QY 744 GCTGACAGTGTTCGTGCGGAAAGTACATTGAGCAATTACTTAGAGACTTTCTGGAGTC 803  
DB 843 GCTGACAGTGTTCGTGCGGAAAGTACATTGAGCAATTACTTAGAGACTTTCTGGAGTC 902  
QY 804 TGCTGACATGTACTTCATCGTTGGCCATCGGGTCATATTTTACGTGATGATAGACACAC 863  
DB 903 TGCTGACATGTACTTCATCGTTGGCCATCGGGTCATATTTTACGTGATGATAGACAC 962  
QY 864 CTCCCGGATGCGCTGTCGACCTGAAACCTCTACATTCCTTACAGTCTTTGAGATCAG 923  
DB 963 CTCCCGGATGCGCTGTCGACCTGAAACCTCTACATTCCTTACAGTCTTTGAGATCAG 1022  
QY 924 GTCTGAGAGAGGTGCGCAGGATATGAGCATGATGCGCATGAGACCATTTGGGAGCACAT 983  
DB 1023 GTCTGAGAGAGGTGCGCAGGATATGAGCATGATGCGCATGAGACCATTTGGGAGCACAT 1082  
QY 984 CTGGCCCCACATCCAGCAGAGTTCGACTTCCTCTTCTGTCATGGAGTGGATCAAGTCTT 1043  
DB 1083 CTGGCCCCACATCCAGCAGAGTTCGACTTCCTCTTCTGTCATGGAGTGGATCAAGTCTT 1142  
QY 1044 TCAAGACAATTCGGGGTGAAACTCTGGGCCAGCTGGTAGACAGCTCCAGGCCTGGTG 1103  
DB 1143 TCAAGACAATTCGGGGTGAAACTCTGGGCCAGCTGGTAGACAGCTCCAGGCCTGGTG 1202  
QY 1104 GTACAAGGCCAGTCCCGAGAAGTTACCTATGAGAGCGGGAACCTCTCGGCCGCTACAT 1163  
DB 1203 GTACAAGGCCAGTCCCGAGAAGTTACCTATGAGAGCGGGAACCTCTCGGCCGCTACAT 1262  
QY 1164 TCATTTCGAGAGGGGATTTTACTACCAACCGCGCATTTTGGAGGACGCTACTCA 1223  
DB 1263 TCATTTCGAGAGGGGATTTTACTACCAACCGCGCATTTTGGAGGACGCTACTCA 1322  
QY 1224 CATTTCTCAACCTCACAGGAGTGTCTTTAAGGGATCTCCAGGACAAAGACATGACAT 1283  
DB 1323 CATTTCTCAACCTCACAGGAGTGTCTTTAAGGGATCTCCAGGACAAAGACATGACAT 1382  
QY 1284 AGAAGCCCAAGTGCATGAGAGCCACCTCAACAATACTTCTTTTCAACAAACCCAC 1343  
DB 1383 AGAAGCCCAAGTGCATGAGAGCCACCTCAACAATACTTCTTTTCAACAAACCCAC 1442  
QY 1344 TAAATTCCTATCTCCAGAGTATTTGCTGGGACTATCAGATAGGCTGCCCTTCAGATATTA 1403  
DB 1443 TAAATTCCTATCTCCAGAGTATTTGCTGGGACTATCAGATAGGCTGCCCTTCAGATATTA 1502  
QY 1404 AAGTGTCAAAGGTAGCTTGGCAGACAAAAGAGTATAATTTGGTTAGAAAATAATCTCGACT 1463

DB 1503 AAGTGTCAAAGGTAGCTTGGCAGACAAAAGAGTATAATTTGGTTAGAAAATAATCTCGACT 1562  
QY 1464 TCAAAATTTGATGGAAACTTGACACTATTCT 1495  
DB 1563 TCAAAATTTGATGGAAACTTGACACTATTACT 1594  
RESULT 10  
US-08-984-900-9  
; Sequence 9, Application US/08984900  
; GENERAL INFORMATION:  
; APPLICANT: d'Apice, Anthony J.F.  
; APPLICANT: Pearce, Martin J.  
; APPLICANT: Robins, Allan J.  
; APPLICANT: Crawford, Robert J.  
; APPLICANT: Rathjen, Peter D.  
; TITLE OF INVENTION: MATERIALS AND METHODS FOR MANAGEMENT OF  
; TITLE OF INVENTION: HYPERACUTE REJECTION IN HUMAN XENOTRANSPLANTATION  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 120 South Sixth Street, Suite 2500  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/984,900  
; FILING DATE:  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/378,617  
; FILING DATE: 26-JAN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ellinger, Mark S.  
; REGISTRATION NUMBER: 34,812  
; REFERENCE/DOCKET NUMBER: 06868/005001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (612) 335-5070  
; TELEFAX: (612) 288-9696  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3450 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
US-08-984-900-9

Query Match 93.9%; Score 1408.4; DB 13; Length 3450;  
Best Local Similarity 96.8%; Pred. No. 0;  
Matches 1483; Conservative 0; Mismatches 11; Indels 38; Gaps 3;  
QY 1 CTTTCCTCTGTAGACTCTTCTTGAATGAGAAGTACCGATTCTGCTGAAGACCTCGCGCT 60  
DB 64 CTTTCCTCTGTAGACTCTTCTTGAATGAGAAGTACCGATTCTGCTGAAGACCTCGCGCT 123  
QY 61 CTCAGGCTCTGGAGTGGAAACCT-CTACCTTCCTTTCTCTGCTGCTGAGCCCTGCTTCT 119  
DB 124 CTCAGGCTCTGGAGTGGAAACCTCTGCTACCTTCCTTTCTGCTGCTGAGCCCTGCTTCT 183  
QY 120 TAGCGAGCCAGAGCTCGACAGAACTCGGTTCCTTTCTGCTGCTTTGGAGGAACACA 179  
DB 184 TCGGCAGCCAGAGCTGACAGAACTCGGTTCCTGCTGCTTTGGAGGAACACA 243  
QY 180 GCTGACATGAGGCTGACTTTTGAACCTCAAGAGATCTGCTTACCCCACTCTCTCTGGAATTA 239

```
Db 244 GCTGACGATGAGGCTGACTTTGAACCTCAAGAGATCTGCTTACCCAGCTCTCCTGGAATTA 303
QY 240 AAGGCGCTGTACTACTATTTGCTGGACCTTAAGATTTTCATCATGATCAGTATGCTTCAAGATCT 299
Db 304 AAGGCGCTGTACTAC-CTTGCTGGACCTAAGATTTTCATCATGATCAGTATGCTTCAAGATCT 362
QY 300 CCATGTCACAAGATCTCCATGTCAGATCCAAAGTCCAAAGTCCAAAGTCCAAAGTCCAAAG 359
Db 363 CCATGTCACAAGATCTCCATGTCAGATCCAAAGTCCAAAGTCCAAAGTCCAAAGTCCAAAG 422
QY 360 ATCTGATCACAGAGAGAAAATATGAATGTCAAGGGGAAAAGTAAATCCTGTTGATGCTGAT 419
Db 423 ATCTGATCACAGAGAGAAAATATGAATGTCAAGGGGAAAAGTAAATCCTGTTGATGCTGAT 482
QY 420 TGTCTCAACCGTGTGCTGCTGTTTGGGAATATGTCAACAG----- 461
Db 483 TGTCTCAACCGTGTGCTGCTGTTTGGGAATATGTCAACAGCCGAGCGCTCTTTCCTT 542
QY 462 -----AATCCAGAGTTGTTGAGAACAGATGCGAGAGGAGCTGGTG 503
Db 543 GTGGATATATCACACAAAAATCCAGAGTTGTTGAGAACAGATGCGAGAGGAGCTGGTG 602
QY 504 GTTCCCAAGCTGTTTAAAAATGGGACCCACAGTTATCAGAGAGACACAGTGAAGGACG 563
Db 603 GTTCCCAAGCTGTTTAAAAATGGGACCCACAGTTATCAGAGAGACACAGTGAAGGACG 662
QY 564 GAGAGAAAAGGTTAGAAATGGAGATCGATTGAAGAGCCCTCAGCTATGGGACTGTTTCAA 623
Db 663 GAGAGAAAAGGTTAGAAATGGAGATCGATTGAAGAGCCCTCAGCTATGGGACTGTTTCAA 722
QY 624 TCCAAAGAACCCCGGATGTTTGTGACAGTGACCCCGTGGAGGCGCCGATGTTGTGGGA 683
Db 723 TCCAAAGAACCCCGGATGTTTGTGACAGTGACCCCGTGGAGGCGCCGATGTTGTGGGA 782
QY 684 AGGACTATTAGACACAGCTGCTGAGAAAGTACTAGCCACACAGAAATCAGTGTGGG 743
Db 783 AGGACTATTAGACACAGCTGCTGAGAAAGTACTAGCCACACAGAAATCAGTGTGGG 842
QY 744 GCTGACAGTGTGCTGTGGAAAGTACATTGAGCATTACTTAGAAGACTTTCTGGAGTC 803
Db 843 GCTGACAGTGTGCTGTGGAAAGTACATTGAGCATTACTTAGAAGACTTTCTGGAGTC 902
QY 804 TGTGATGATGATCTGATGTTGGCCATCGGGTCATATTTTTCATGATAGACGACAC 863
Db 903 TGTGATGATGATCTGATGTTGGCCATCGGGTCATATTTTTCATGATAGATGACAC 962
QY 864 CTCCCGGATGCTGCTGACACCTGACCTCTACATTCCTTACAGTCTTTGAGATCAG 923
Db 963 CTCCCGGATGCTGCTGACACCTGACCTCTACATTCCTTACAGTCTTTGAGATCAG 1022
QY 924 GTCTGAGAAGAGGTGGCAGGATATCAGCATGATGCGCATGAAGACCATTTGGGAGGACAT 983
Db 1023 GTCTGAGAAGAGGTGGCAGGATATCAGCATGATGCGCATGAAGACCATTTGGGAGGACAT 1082
QY 984 CTGGGCCCCATCCAGACAGAGTGCAGCTTCTCTTCTGATGAGAGTGAATCAAGTCTT 1043
Db 1083 CTGGGCCCCATCCAGACAGAGTGCAGCTTCTCTTCTGATGAGAGTGAATCAAGTCTT 1142
QY 1044 TCAAGACAACTTCGGGGTGAACCTCTGGGCCAGCTGGTAGCACAGCTCCAGGCGTGGTG 1103
Db 1143 TCAAGACAACTTCGGGGTGAACCTCTGGGCCAGCTGGTAGCACAGCTCCAGGCGTGGTG 1202
QY 1104 GTACAAAGGCGAGTCCCGAGAAAGTTACCTATGAGAGCGGGAACTGTGCGCGCGGTACAT 1163
Db 1203 GTACAAAGGCGAGTCCCGAGAAAGTTACCTATGAGAGCGGGAACTGTGCGCGCGGTACAT 1262
QY 1164 TCCATTGGGAGAGGGGATTTTCTACTACACGCGGCGCATTTTGGAGGAACGCGCTACTCA 1223
Db 1263 TCCATTGGGAGAGGGGATTTTCTACTACACGCGGCGCATTTTGGAGGAACGCGCTACTCA 1322
QY 1224 CATTCTCAACCTCAGGAGGAGTCTTTAAGGGGATCCTCCAGGACAGAAAGATGACAT 1283
|||||
```

```
Db 1323 CATTCTCAACCTCAGGAGGAGTCTTTAAGGGGATCCTCCAGGACAGAAACATGACAT 1382
QY 1284 AGAAGCCAGTGGGATGATGAGAGCCACCTCAACAAATACTTCTCTTCAACAAACCCAC 1343
Db 1383 AGAAGCCAGTGGGATGATGAGAGCCACCTCAACAAATACTTCTCTTCAACAAACCCAC 1442
QY 1344 TAAATCCTATCTCCAGAGTATGCTGGGACTATCAGATAGGCTTGCCTTCAGATATTTAA 1403
Db 1443 TAAATCCTATCTCCAGAGTATGCTGGGACTATCAGATAGGCTTGCCTTCAGATATTTAA 1502
QY 1404 AAGTGTCAAGGTAGCTTGGGACAGACAAAAGAGTAAATTTGGTTAGAAATATGCTGACT 1463
Db 1503 AAGTGTCAAGGTAGCTTGGGACAGACAAAAGAGTAAATTTGGTTAGAAATATGCTGACT 1562
QY 1464 TCAAAATGTGATGAAACTTGCACACTATTTCT 1495
Db 1563 TCAAAATGTGATGAAACTTGCACACTATTTACT 1594

RESULT 11
US-08-213-200A-3
: Sequence 3, Application US/08213200A
: GENERAL INFORMATION:
: APPLICANT: Galili, Dr. Uri
: TITLE OF INVENTION: REPLIC, DR. PATRICIA M.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR VACCINES
: TITLE OF INVENTION: COMPRISING ALPHA-GALACTOSYL EPITOPES
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SEIDEL, GONDA, LAVORNA & MONACO, P.C.
: STREET: Suite 1800, Two Penn Center Plaza
: CITY: Philadelphia
: STATE: PA
: COUNTRY: USA
: ZIP: 19102
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/213, 200A
: FILING DATE: 15-MAR-1994
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Daniel A. Monaco
: REGISTRATION NUMBER: 30,480
: REFERENCE/DOCKET NUMBER: 8760-2 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (215) 568-8383
: TELEFAX: (215) 568-5549
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1131 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cdna
: US-08-213-200A-3
```

Query Match 67.1%; Score 1005.8; DB 6; Length 1131;  
Best Local Similarity 95.3%; Pred. No. 9e-284;  
Matches 1078; Conservative 0; Mismatches 2; Indels 51; Gaps 2;

```
QY 382 ATGAATGTCAGGGAAGTAATCTGTTGATGCTGATTTGCTCAACCGTGGTTGCTG 441
Db 1 ATGAATGTCAGGGAAGTAATCTGTTGATGCTGATTTGCTCAACCGTGGTTGCTG 60
QY 442 TTTTGGGAATATGTCACAG-----AATT 465
Db 61 TTTTGGGAATATGTCACAGCCCCAGAGCGCTCTTCTTGTGGATATATCACACAAAATT 120
|||||
```

```
QY 466 CCAGAGGTTGGTGAGAACAGATGGCAGAGAGACTGGTGGTTCCTCCCAAGCTGGTTTAAAAAT 525
Db 121 CCAGAGGTTGGTGAGAACAGATGGCAGAGAGACTGGTGGTTCCTCCCAAGCTGGTTTAAAAAT 180
QY 526 GGGACCCACAGTTATCAAGAACACACACCTAGAGAGCAGAGAGAAAGGGTAGAAATGGA 585
Db 181 GGGACCCACAGTTATCAAGAACACACACCTAGAGAGCAGAGAGAAAGGGTAGAAATGGA 240
QY 586 GATCGCATTTGA-----AGAGCCTCAGCTATGGGACTATGGGACTGGTTCAATCCAAAG 630
Db 241 GATCGCATTTGAGGAAGATCACACACAGAGCCTCAGCTATGGGACTGGTTCAATCCAAAG 300
QY 631 AACCGCCCGGATGTTTGTACAGTACCCCGTGGAGAGCCCGGATTTGTGTGGGAAGCACT 690
Db 301 AACCGCCCGGATGTTTGTACAGTACCCCGTGGAGAGCCCGGATTTGTGTGGGAAGCACT 360
QY 691 TATGACACAGCTCTGCTGGAAGTACTACGCCACACAGAAACTCACTCTGGGGCTGACA 750
Db 361 TATGACACAGCTCTGCTGGAAGTACTACGCCACACAGAAACTCACTCTGGGGCTGACA 420
QY 751 GTGTTTGTCTGTGGGAAGTACATTGAGCATTACTTTAGAGACTTTCTGGAGTCTGCTGAC 810
Db 421 GTGTTTGTCTGTGGGAAGTACATTGAGCATTACTTTAGAGACTTTCTGGAGTCTGCTGAC 480
QY 811 ATGTACTTTCATGGTTGGCCATCGGGTTCATATTTTACGTGATGATAGACACACCTCCCGG 870
Db 481 ATGTACTTTCATGGTTGGCCATCGGGTTCATATTTTACGTGATGATAGACACACCTCCCGG 540
QY 871 ATGCCCTGTGCTGACCTGAACCTCTACATTCTTACAAGTCTTTGAGATCAGGTCGTAG 930
Db 541 ATGCCCTGTGCTGACCTGAACCTCTACATTCTTACAAGTCTTTGAGATCAGGTCGTAG 600
QY 931 AAGAGTGGCAGGATATCAGCATGATGCCATGAAGACATTGGGGAGCACATCTCTGGCC 990
Db 601 AAGAGTGGCAGGATATCAGCATGATGCCATGAAGACATTGGGGAGCACATCTCTGGCC 560
QY 991 CACATCCACAGCAGGTGCGACTTCCTCTCTCTGATGAGAGTGGATCAAGTCTTTCAAGAC 1050
Db 661 CACATCCACAGCAGGTGCGACTTCCTCTCTCTGATGAGAGTGGATCAAGTCTTTCAAGAC 720
QY 1051 AACTTCGGGGTGGAACTCTGGGAGCTGGTAGCAGCTCCAGCCTCGTGGTACAAG 1110
Db 721 AACTTCGGGGTGGAACTCTGGGAGCTGGTAGCAGCTCCAGCCTCGTGGTACAAG 780
QY 1111 GCCAGTCCCGAGAAGTTACCTATGAGAGCGGGAACCTGTCGGCGCGTACATTCATTTC 1170
Db 781 GCCAGTCCCGAGAAGTTACCTATGAGAGCGGGAACCTGTCGGCGCGTACATTCATTTC 840
QY 1171 GGAGAGGGGATTTTACTACACAGCGGCGCATTTTGGAGGAGCAGCTACTCACATTCCTC 1230
Db 841 GGAGAGGGGATTTTACTACACAGCGGCGCATTTTGGAGGAGCAGCTACTCACATTCCTC 900
QY 1231 AACCTCACCAGGAGTGCTTTAAGGGATCTCCAGGACAAGAACATCACATAGAAGCC 1290
Db 901 AACCTCACCAGGAGTGCTTTAAGGGATCTCCAGGACAAGAACATCACATAGAAGCC 960
QY 1291 CAGTGGCATGATGAGAGCCACCTCAACAAATACTTCTCTTTTCAACAAACCCACTAAAATC 1350
Db 961 CAGTGGCATGATGAGAGCCACCTCAACAAATACTTCTCTTTTCAACAAACCCACTAAAATC 1020
QY 1351 CTATCTCCAGATATTGCTGGGACTATCAGATAGGCGCTGCCCTTCAGATATTAAAAGTGTG 1410
Db 1021 CTATCTCCAGATATTGCTGGGACTATCAGATAGGCGCTGCCCTTCAGATATTAAAAGTGTG 1080
QY 1411 AAGGTAGCTTGGCAGACAAAGATATAATTTGGTTAGAAATAATGCTCTGA 1461
Db 1081 AAGGTAGCTTGGCAGACAAAGATATAATTTGGTTAGAAATAATGCTCTGA 1131
```

RESULT 12

US-09-173-270-3

: Sequence 3, Application US/09173270

: GENERAL INFORMATION:

```
;
;
; APPLICANT: GALILI, URI
; REPIK, PATRICIA M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR VACCINES
; COMPRISING ALPHA-GALACTOSYL EPITOPES
;
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.
; STREET: Suite 1800, Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/173,270
; FILING DATE: 23-Mar-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/704,548
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8760-2 C11
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1131 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-173-270-3
```

```
Query Match 67.1%; Score 1005.8; DB 15; Length 1131;
Best Local Similarity 95.3%; Pred. No. 9e-284;
Matches 1078; Conservative 0; Mismatches 2; Indels 51; Gaps 2;

QY 382 ATGAATGTCAAGGGAAGAACTAATCTCTGTTGATGCTGCTCAACCCGTGGTTCGCTG 441
Db 1 ATGAATGTCAAGGGAAGAACTGATCTCTGTTGATGCTGCTCAACCCGTGGTTCGCTG 60
QY 442 TTTTGGGAATATGTCAACAG-----AATTT 465
Db 61 TTTTGGGAATATGTCAACAGCCCGAGACGCTCTTCTTCTGGATATATCACACAAATTT 120
QY 466 CCAGAGGTTGGTGAGAACAGATGGCAGAGAGACTGGTGGTTCCTCCCAAGCTGGTTTAAAAAT 525
Db 121 CCAGAGGTTGGTGAGAACAGATGGCAGAGAGACTGGTGGTTCCTCCCAAGCTGGTTTAAAAAT 180
QY 526 GGGACCCACAGTTATCAAGAACACACACCTAGAGAGCAGAGAGAAAGGGTAGAAATGGA 585
Db 181 GGGACCCACAGTTATCAAGAACACACACCTAGAGAGCAGAGAGAAAGGGTAGAAATGGA 240
QY 586 GATCGCATTTGA-----AGAGCCTCAGCTATGGGACTGGTTCAATCCAAAG 630
Db 241 GATCGCATTTGAGGAAGATCACACACAGAGCCTCAGCTATGGGACTGGTTCAATCCAAAG 300
QY 631 AACCGCCCGGATGTTTGTACAGTACCCCGTGGAGAGCCCGGATTTGTGTGGGAAGCACT 690
Db 301 AACCGCCCGGATGTTTGTACAGTACCCCGTGGAGAGCCCGGATTTGTGTGGGAAGCACT 360
QY 691 TATGACACAGCTCTGCTGGAAGTACTACGCCACACAGAAACTCACTCTGGGGCTGACA 750
Db 361 TATGACACAGCTCTGCTGGAAGTACTACGCCACACAGAAACTCACTCTGGGGCTGACA 420
```



Qy	751	GTGTTTTCCTGTCGGAAGTACATTGAGCATTTACTTAGAAGACTTTCTGGAGTCTGCTGAC	810
Db	421	GTGTTTTCCTGTCGGAAGTACATTGAGCATTTACTTAGAAGACTTTCTGGAGTCTGCTGAC	480
Qy	811	ATGTACTTTCATGTTGGCCATCGGGTCATATTTTACGTCAATGAGCAGCACCTCCCGG	870
Db	481	ATGTACTTTCATGTTGGCCATCGGGTCATATTTTACGTCAATGAGCAGCACCTCCCGG	540
Qy	871	ATGCCTGTCTGCACCTGAACCTCTACATTTCTTTACAAGTCTTTTGAGATCAGGTCTGAG	930
Db	541	ATGCCTGTCTGCACCTGAACCTCTACATTTCTTTACAAGTCTTTTGAGATCAGGTCTGAG	600
Qy	931	AAGAGTGGCAGGATATCAGCATGATCGGCATGAAGACCATTTGGGAGCACATCCTGGCC	990
Db	601	AAGAGTGGCAGGATATCAGCATGATCGGCATGAAGACCATTTGGGAGCACATCCTGGCC	660
Qy	991	CACATCCAGCACGAGTCCGACTTCTCTTCTGTCATGGAGCTGGATCAAGTCTTTTCAAGAC	1050
Db	661	CACATCCAGCACGAGTCCGACTTCTCTTCTGTCATGGAGCTGGATCAAGTCTTTTCAAGAC	720
Qy	1051	AAC TTCGGGTGGAAACTCTGGGCCAGCTGGTAGCACAGCTCCAGGCCCTGGTGTACAAG	1110
Db	721	AAC TTCGGGTGGAAACTCTGGGCCAGCTGGTAGCACAGCTCCAGGCCCTGGTGTACAAG	780
Qy	1111	GCAGTCCCAGAAAGTTTCACTATGAGAGCGCGGAAC TGTGCGGCCGTACATTCATTC	1170
Db	781	GCAGTCCCAGAAAGTTTCACTATGAGAGCGCGGAAC TGTGCGGCCGTACATTCATTC	840
Qy	1171	GGAGAGGGGATTTTACTACACGCGGCCATTTTGGAGGAACGCCTACTACATTTCTC	1230
Db	841	GGAGAGGGGATTTTACTACACGCGGCCATTTTGGAGGAACGCCTACTACATTTCTC	900
Qy	1231	AAC TCACAGGAGGTGCTTTTAAGGGATTCCTCCAGGACAAGAAACATGACATAGAACC	1290
Db	901	AAC TCACAGGAGGTGCTTTAAGGGATTCCTCCAGGACAAGAAACATGACATAGAACC	960
Qy	1291	CAGTGGCATGATCAGAGCCACCTCAACAATACTTCTTTTCAACAACCCACTAAATC	1350
Db	961	CAGTGGCATGATCAGAGCCACCTCAACAATACTTCTTTTCAACAACCCACTAAATC	1020
Qy	1351	CTATCTCCAGAGTATTTGCTGGGACTATCAGATAGGCCTGCCCTTCAGATATTTAAAAGTCTC	1410
Db	1021	CTATCTCCAGAGTATTTGCTGGGACTATCAGATAGGCCTGCCCTTCAGATATTTAAAAGTCTC	1080
Qy	1411	AAGTAGCTTTGGCAGACAAAGAGTATAATTTGGTTAGAAATAATGCTGA	1461
Db	1081	AAGTAGCTTTGGCAGACAAAGAGTATAATTTGGTTAGAAATAATGCTGTA	1131

RESULT 13

US-08-214-580-2

```

: Sequence 2, Application US/08214580
:
: GENERAL INFORMATION:
:
: APPLICANT: Sandrin, Mauro S.
: APPLICANT: McKenzie, Ian F. C.
: TITLE OF INVENTION: XENOTRANSPLANTATION
: TITLE OF INVENTION: THERAPIES
:
: NUMBER OF SEQUENCES: 4
:
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Maurice W. Klee
: STREET: 1951 Burr Street
: CITY: Fairfield
: STATE: Connecticut
: COUNTRY: USA
: ZIP: 06430
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3 1/2 inch, 750 kb storage
:
: COMPUTER: Dell 486/50
: OPERATING SYSTEM: DOS 6.2
: SOFTWARE: WordPerfect 6.0
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/214,580

```

QY 951 CATGATCGCGCATGAAGACCATTTGGGAGACACATCTCGGCCACATCCAGCACGAGTCTGA 1010  
Db 619 CATGATCGCGCATGAAGACCATCGGGAGACACATCTCGGCCACATCCAGCACGAGTCTGA 678  
QY 1011 CTTCCTCTCTCGATGACGCTGATCAAGTCTTTCAAGACAACCTCGGGGTGAAACTCT 1070  
Db 679 CTTCCTCTCTCGATGACGCTGATCAAGTCTTTCCAAACAACCTTTGGGTGGAGACCT 738  
QY 1071 GGCCAGCTGTTAGCACAGCTCCAGGCGTGGTGTCAAGGCCAGTCCCGAGAAGTTTCA 1130  
Db 739 GGCCAGCTGTTAGCACAGCTCCAGGCGTGGTGTCAAGGCCAGTCCCGAGAAGTTTCA 798  
QY 1131 CTATGAGAGCGGGAACCTGTGCGCGCGGTACATTCATTCGGAGAGGGGATTTTACTA 1190  
Db 799 CTACGAGAGCGGGAAGGAGTCCGAGCTACATTCGCTTTGGCCAGGGGATTTTATTA 858  
QY 1191 CCAGCGGGCCATTTTGGAGGAAGCGCTACTACATCTCAACCTCACAGGAGTCTT 1250  
Db 859 CCAGCGAGCCATTTTGGGGAACACCCACTCAGGTTCTAAACATCACTCAGGAGTCTT 918  
QY 1251 TAAGGGGATCCTCCAGGACAAGAAACATGACATAGAACCCAGTGGCGATGATGAGAGCCA 1310  
Db 919 CAAGGAATCCTCCAGGACAAGAAATGACATAGAACCGAGTGGCATGATGAAGCCA 978  
QY 1311 CCTCAACAATACTCTCTTTTCAACAAACCCACTTAAATCTATCTCCAGAGTATTGCTG 1370  
Db 979 TCTAAACAAGTATTTCTCTTCAACAAACCCACTTAAATCTTATCCCAAGTACTGCTG 1038  
QY 1371 GGACTATCAGATAGGCTGCTTCAGATATTAAGTGTCAAGTAGCTTTGGCAGACAAA 1430  
Db 1039 GGATATATCATATAGGATGCTGTGGATATTAGGATTTCAAGATAGCTTTGGCAGAAAA 1098  
QY 1431 AGAGTAAATTTGGTTAGAAATTAATGCTGACTTCAAAATTTGTATGGAAACTT 1483  
Db 1099 AGAGTAAATTTGGTTAGAAATTAACATCTGACTTTAAATTTGTCCAGCAGCTT 1151

RESULT 14

US-08-260-201-5  
; Sequence 5, Application us/08260201  
; GENERAL INFORMATION:  
; APPLICANT: Sandrin, Mauro S.  
; APPLICANT: Fodor, William L.  
; APPLICANT: Squinto, Stephen P.  
; APPLICANT: McKenzie, Ian F. C.  
; TITLE OF INVENTION: Methods for Reducing  
; TITLE OF INVENTION: Hyperacute Rejection of Xenografts  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Maurice M. Klee  
; STREET: 1951 Burr Street  
; CITY: Fairfield  
; STATE: Connecticut  
; COUNTRY: USA  
; ZIP: 06430  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 Inch, 750 kb storage  
; COMPUTER: Dell 486/50  
; OPERATING SYSTEM: DOS 6.2  
; SOFTWARE: WordPerfect 6.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/260, 201  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION NUMBER:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Klee, Maurice M.  
; REGISTRATION NUMBER: 30,399  
; REFERENCE/DOCKET NUMBER: ALX-144  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (203) 255-1400

; TELEFAX: (203) 254-1101  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1423 base pairs  
; TYPE: Nucleic Acid  
; STRANDEDNESS: Double  
; TOPOLOGY: Linear  
; MOLECULE TYPE: cDNA to mRNA  
; DESCRIPTION: galactosyl transferase.  
; DESCRIPTION: full coding sequence  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Sus scrofa  
; US-08-260-201-5

Query Match 48.4%; Score 726.6; DB 6; Length 1423;  
Best Local Similarity 78.8%; Pred. No. 1.le-201;  
Matches 893; Conservative 0; Mismatches 234; Indels 6; Gaps 2;  
QY 354 CTCAAGATCTGGATCACAGGAGAAATAATGAATGTCAAGGGAAACATAATCTCTGTTGAT 413  
Db 22 CCAGCTTCTGCCGATCAGGAGAAATAATGAATGTCAAGAGAGAGGTGTTCTGTCTCAT 81  
QY 414 GCTGATGTTCTCAACCGTGTGTTGTTGGGAATATCTCAACAGAAATCTCCAGAGGT 473  
Db 82 GCTGCTGTCTCAACTGTAATGTTGTTGGGAATACATCAACAGAAACCCAGAAGT 141  
QY 474 TGTGAGAACAGATGGGAGAGGACTGTTGTTCCCAAGCTGTTTAAAATGGGAGCCA 533  
Db 142 TGG---CAGCAGTGTCTCAGAGGGGCTGTTGTTCCGAGGTGTTTAAACAATGGGATCA 198  
QY 534 CAGTTATCAAGAAGACAACGCTAGAACGCGAGAGAAAGS---GTAGAAATGGAGATCG 590  
Db 199 CAGTTACCAGAAAGAAAGACCGCTATAGCAAGAAAGAAAGAAAGAAAGAAAGACAA 258  
QY 591 CATTGAAGAGCCTCAGCTATGGGACTGGTTCAATCCAAAGAACCGCCCGATGTTTGTAC 650  
Db 259 CAGAGGAGAGCTTCCGCTAGTGGACTGGTTTAATCTCTGAGAAACGCCAGAGTCTGTGAC 318  
QY 651 AGTCACCCCGTGAAGCGCCGATTTGTTGGGAAGGCACTTATGACACACGCTGCTGCTGA 710  
Db 319 CATTAACCCAGATGGAAGGCTCCAGTGGTATGGGAAGGCACTTACAAACAGAGCCGCTTAGA 378  
QY 711 AAAGTACTACGACACACAGAAATCACTGTGGGCTCAGAGTGTGTTGCTGTGGGAAAGTA 770  
Db 379 TAATTATTATGCCAAACAGAAATATCCGTTGGGCTTACCGGTTTCTGCTGTGCGAAGATA 438  
QY 771 CATTGAGCATTTACTTGAAGACTTTTCTGGAGTCTGTGACATGCTACTTCTATGTTGGCCA 830  
Db 439 CATTGAGCATTTACTTGGAGGAGTCTTAAATATCTGCAAAATACATCTTCAATGTTGGCCA 498  
QY 831 TCGGGTCATATTTTACGTCATGATAGACGACACCTCCCGGATGCCCTGCTGTCACCTGAA 890  
Db 499 CAAAGTCATCTTTTACATCATGTTGGTATGATATCTCCAGGATGCCCTTGTATAGAGTGG 558  
QY 891 CCCTCTACATTTCTTACAAGTCTTTGAGATCAGGCTCTGAGAAGAGTGGCAGAGATATCAG 950  
Db 559 TCCTCTCGTTCCTTTTAAAGTGTTTGAGATCAAGTCCGAGAAGAGTGGCAAGACATCAG 618  
QY 951 CATGATCGGCATGAAGACCATTTGGGAGACACATCTCTGCGCCACATCCAGCACGAGGTCTGA 1010  
Db 619 CATGATCGGCATGAAGACCATCGGGAGACACATCTCTGCGCCACATCCAGCACGAGGTCTGA 678  
QY 1011 CTTCCTCTTCTGATGAGGATGATCAAGTCTTTTCAAGACAACCTTCGGGTGGAAACTCT 1070  
Db 679 CTTCCTCTTCTGATGAGGATGATCAAGTCTTTTCAAGACAACCTTTTGGGTGGAGACCT 738  
QY 1071 GGCCAGCTGTAGCACAGCTCCAGGCGTGGTGTGTAAGGCCAGTCCCGAGAAGTCTCA 1130  
Db 739 GGCCAGCTGTAGCACAGCTCCAGGCGTGGTGTGTAAGGCCAGTCTTCACAGAGTCTCA 798

Qy 1131 CTATGAGAGCGGGAAGTCTCGCGCGGTACATTCATTCGGAGAGGGGAGTCTTACTA 1190  
Db 799 CTACGAGAGCGGGAAGAGTCCGCGAGCTACATTCCTGTTGGCCAGGGGAGTCTTATTA 858  
Qy 1191 CCAGCGCGCATTTTGGAGGAAGCGCTACTACATTCCTCAACCTCACCAGGGAGTCTT 1250  
Db 859 CCAGCGAGCATTTTGGGGAAGACCCACTCAGGTCTTAAACATCACTCAGGAGTCTT 918  
Qy 1251 TAAGGGATCTCCAGGACAAGAACATGACATAGAAGCCAGTGGCATGATGAGAGCCA 1310  
Db 919 CAAGGGAATCCTCAGGACAGGAATGACATAGAAGCCAGTGGCATGATGAGAGCCA 978  
Qy 1311 CCTCAACAAATCTCTCTTTTCAACAAACCCACTAAATCTCTCAGAGTATCTGTG 1370  
Db 979 TCTAAACAAGTATTTCTCTTCTCAACAAACCCACTAAATCTTATCCCCAGAACTGCTG 1038  
Qy 1371 GGACTATCAGATAGCGCTGCTTCAAGATATTAAGTCTCAAGCTAGCTGGCAGACAAA 1430  
Db 1039 GGATTATCATATAGCATGTCTGGGATATTTAGGATTTCTCAAGATAGCTTGGCAGAAAA 1098  
Qy 1431 AGATATAAATTTGGTTAGAAATATGTCTGACTTCAAAATTTGTATGGAAACTT 1483  
Db 1099 AGATATAAATTTGGTTAGAAATATCTGACTTTAAATTTGTCGACAGTCTT 1151

RESULT 15

US-08-278-282-5  
; Sequence 5, Application US/08278282  
; GENERAL INFORMATION:  
; APPLICANT: Sandrin, Mauro S.  
; APPLICANT: Fodor, William L.  
; APPLICANT: Rother, Russell P.  
; APPLICANT: Squinto, Stephen P.  
; APPLICANT: McKenzie, Ian F. C.  
; TITLE OF INVENTION: Methods for Reducing  
; TITLE OF INVENTION: Hyperacute Rejection of Xenografts  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Maurice M. Klee  
; STREET: 1951 Burr Street  
; CITY: Fairfield  
; STATE: Connecticut  
; COUNTRY: USA  
; ZIP: 06430  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 750 Kb storage  
; COMPUTER: Dell 486/50  
; OPERATING SYSTEM: DOS 6.2  
; SOFTWARE: WordPerfect 6.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/278,282  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/260,201  
; FILING DATE: June 15, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Klee, Maurice M.  
; REGISTRATION NUMBER: 30,399  
; REFERENCE/DOCKET NUMBER: ALX-144.1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (203) 255-1400  
; TELEFAX: (203) 254-1101  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1423 base pairs  
; TYPE: Nucleic Acid  
; STRANDEDNESS: Double  
; TOPOLOGY: Linear  
; MOLECULE TYPE: cDNA to mRNA  
; DESCRIPTION: galactosyl transferase,  
; DESCRIPTION: full coding sequence  
; HYPOTHETICAL: No

; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Sus scrofa  
US-08-278-282-5  
  
Query Match 48.4%; Score 726.6; DB 6; Length 1423;  
Best Local Similarity 78.8%; Pred. No. 1.1e-201;  
Matches 893; Conservative 0; Mismatches 234; Indels 6; Gaps 2;  
  
Qy 354 CTCAAGATCTGGATCAGAGGAGAAATTAATGTCAGGAGAAAGTAATCTCTGTTGAT 413  
Db 22 CCCAGCTTCTGCCGATCAGAGAGAAATTAATGAATGTCAGAGGAGAGTGGTCTGTCTCAAT 81  
Qy 414 GCTGATTTCTCAACCGTGGTGTCTGGGAATATCTCAACAGAAATTCAGAGGT 473  
Db 82 GCTGCTTCTCACTGTAATGGTGTGGGAATACATCAACAGAAACCCAGAGT 141  
Qy 474 TGGTGAGAACAGATGGCAGAGAGTGGTGTCTCCAGAGTGGTAAATGGAGCCCA 533  
Db 142 TGG---CAGCAGTCTCAGAGGGCTGGTGTCTCCAGAGTGGTAAATGGAGTCA 198  
Qy 534 CAGTTATCAAGACACACACCTAGAGGACGAGAGAGAGG---GTAGAAATGGAGATCG 590  
Db 199 CAGTTACACGAGAGAGAGAGCTATAGGCAACGAGAAAGAAAGAAAGAGAGCA 258  
Qy 591 CATTGAAGAGCTCAGCTATGGGACTGTGTTCAATCCAAAGAACCGCCGGATGTTTGAC 650  
Db 259 CAGAGGAGAGCTCCGCTAGTGGACTGTTTAACTCAGAGAACCCAGAGTGGTGCAC 318  
Qy 651 AGTGACCCGCTGGAAGGCGCGATTTGTGGGAAGGCACTTATGACACAGCTCTGCTGGA 710  
Db 319 CATAAACAGATGGAAGGCTCCAGTGGTATGGGAAGGCACTTACAACAGAGCGCTTAGA 378  
Qy 711 AAAGTACTACCCACACAGAAACTCAGTGTGGGCTGACAGTGTCTGCTGGGAAAGTA 770  
Db 379 TAATTTATGCGCAACAGAAAAATACCGTGGGCTTGACGTTTGGTCTGCGGAAGATA 438  
Qy 771 CATTGAGCATTTAGAAAGCTTTCTGGAGTCTGCTGACATGTACTCATCTGTTGGGCA 830  
Db 439 CATTGAGCATTTAGGAGGATTTTAAATATCTGCAATACATATCTCATCTGTTGGGCA 498  
Qy 831 TCGGGTCATATTTACGTGATGATGAGGACACCTCCCGGATGCTGTGCTGCACTGAA 890  
Db 499 CAAAGTCACTCTTTACATCATGCTGATGATATCTCCAGAGTGGCTTTGATAGAGTGG 558  
Qy 891 CCCTCTACATTCCTTACAAGCTTTTGAGATCAGGCTGAGAGAGGCTGGCAGGATATCAG 950  
Db 559 TCCTCTCGTTTCTTTAAAGTGTGTTGAGATCAAGTCCGAGAGAGGTTGGCAAGACATCAG 618  
Qy 951 CATGATCGCATGAAGACCATTTGGGGAGCACATCTTGCCCCACATCCAGCAGAGGTGGA 1010  
Db 619 CATGATCGCATGAAGACCATCGGGAGCACATCTTGCCCCACATCCAGCAGAGGTGGA 678  
Qy 1011 CTTCTCTCTGATGGAGTGGATCAAGTCTTTCAAGACAACCTTCGGGGTGGAACTCT 1070  
Db 679 CTTCTCTCTGATGACGTTGGATCAGGCTGATGATATCTCCAAACAACTTTGGGGTGGAGACCT 738  
Qy 1071 GGGCAGCTGCTAGCACAGCTCCAGGCTGGTGTGTAAGGCGAGTCCCGAGAAAGTTTCA 1130  
Db 739 GGGCAGCTGCTGCTAGCTAGAGGCTGGTGTGTAAGGCGACATCTGACAGAGTTCAC 798  
Qy 1131 CTATGAGAGCGGGAACCTGTCGGCGGCTACATTTCCATTCGAGAGGGGATTTTTACTA 1190  
Db 799 CTACGAGAGCGGGAAGGAGTCCGACAGCTACATTTCCGTTTGGCCAGGGGATTTTTATTA 858  
Qy 1191 CCAGCGGCGCATTTTGGAGGAAGCGCTACTTCAATCTCAACCTCACCAGGGAGTCTT 1250  
Db 859 CCAGCGAGCATTTTGGGGAAGACCCACTCAGGTCTTAAACATCACTCAGGAGTGGT 918  
Qy 1251 TAAGGGATCTCTCAGGACAAAGAAATGACATAGAAGCCAGTGGCATGATGAGAGCCA 1310  
Db 919 CAAGGGAATCTCCAGGACAAAGAAATGACATAGAAGCCAGTGGCATGATGAGAGCCA 978

QY 1311 CCTCAACAATACTTCCTTTTCAACAAACCCACTAAAAATCCTATCTCCAGAGTATTGCTG 1370  
  ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 979 TCTAAACAAGTATTTCCTTCTCAACAACCCACTAAAAATCTTATCCCCAGAACTACTGCTG 1038  
  ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 1371 GGACTATCAGATAGGCTGCCCTTCAGATATTAAAAAGTGTCAAGGTAGCTTGGCAGACAAA 1430  
  ||||| ||||| ||||| || || ||||| || ||||| || ||||| ||||| ||||| |||||  
Db 1039 GGATTATCATATAGGCATGCTGTGGATATTAGGATTGTCAAGATAGCTTGGCAGAAAAA 1098  
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 1431 AGAGTATAATTTGGTTAGAAAATAATGCTGACTTCAAAATTGTCATGCAAACTT 1483  
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1099 AGAGTATAATTTGGTTAGAAAATAACATCTGACTTTAAATTGTGCCAGAGTTT 1151  
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Search completed: May 17, 2002, 20:30:28  
Job time: 9097 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2002, 11:05:57 ; Search time 109.08 Seconds  
(without alignments)  
1268.739 Million cell updates/sec

Title: US-09-863-475A-4  
Perfect score: 2136  
Sequence: 1 MITMLQDLHVNKISMSRSKS.....IKSVKVAQOTKEYNLVRNVV 394

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3516493 seqs, 351254056 residues  
Total number of hits satisfying chosen parameters: 3516493

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*

1:	/cgn2_6/ptodata/2/paa/pctus_COMB.pep.*
2:	/cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3:	/cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4:	/cgn2_6/ptodata/2/paa/US080_COMB.pep.*
5:	/cgn2_6/ptodata/2/paa/US081_COMB.pep.*
6:	/cgn2_6/ptodata/2/paa/US082_COMB.pep.*
7:	/cgn2_6/ptodata/2/paa/US083_COMB.pep.*
8:	/cgn2_6/ptodata/2/paa/US084_COMB.pep.*
9:	/cgn2_6/ptodata/2/paa/US085_COMB.pep.*
10:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*
11:	/cgn2_6/ptodata/2/paa/US087_COMB.pep.*
12:	/cgn2_6/ptodata/2/paa/US088_COMB.pep.*
13:	/cgn2_6/ptodata/2/paa/US089_COMB.pep.*
14:	/cgn2_6/ptodata/2/paa/US090_COMB.pep.*
15:	/cgn2_6/ptodata/2/paa/US091_COMB.pep.*
16:	/cgn2_6/ptodata/2/paa/US092_COMB.pep.*
17:	/cgn2_6/ptodata/2/paa/US093_COMB.pep.*
18:	/cgn2_6/ptodata/2/paa/US094_COMB.pep.*
19:	/cgn2_6/ptodata/2/paa/US095_COMB.pep.*
20:	/cgn2_6/ptodata/2/paa/US096_COMB.pep.*
21:	/cgn2_6/ptodata/2/paa/US097_COMB.pep.*
22:	/cgn2_6/ptodata/2/paa/US098_COMB.pep.*
23:	/cgn2_6/ptodata/2/paa/US099_COMB.pep.*
24:	/cgn2_6/ptodata/2/paa/US60_COMB.pep.*
25:	/cgn2_6/ptodata/2/paa/US60_MERGED_COMB.pep1.*
26:	/cgn2_6/ptodata/2/paa/US10_MERGED_COMB.pep1.*
27:	/cgn2_6/ptodata/2/paa/US09_MERGED_COMB.pep2.*
28:	/cgn2_6/ptodata/2/paa/US09_MERGED_COMB.pep1.*
29:	/cgn2_6/ptodata/2/paa/US08_MERGED_COMB.pep1.*
30:	/cgn2_6/ptodata/2/paa/US08_MERGED_COMB.pep1.*
31:	/cgn2_6/ptodata/2/paa/US06_MERGED_COMB.pep1.*
32:	/cgn2_6/ptodata/2/paa/PCT_MERGED_COMB.pep1.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	2136	100.0	394	7 US-08-379-040-2 Sequence 2, Appl1

2	2136	100.0	394	12 US-08-823-489-4 Sequence 4, Appl1
3	2136	100.0	394	19 US-09-593-316-10 Sequence 10, Appl1
4	2136	100.0	394	22 US-09-863-475-4 Sequence 4, Appl1
5	2136	100.0	394	27 US-09-863-475A-4 Sequence 4, Appl1
6	2136	100.0	394	27 US-09-946-034-2 Sequence 2, Appl1
7	1969	92.2	359	28 US-09-994-427A-11 Sequence 11, Appl1
8	1950	91.3	371	5 US-08-188-607B-12 Sequence 12, Appl1
9	1950	91.3	371	7 US-08-378-617-12 Sequence 12, Appl1
10	1950	91.3	371	13 US-08-984-900-12 Sequence 14, Appl1
11	1950	91.3	371	28 US-09-870-759-74 Sequence 74, Appl1
12	1542.5	72.2	376	28 US-09-994-427A-12 Sequence 12, Appl1
13	1539.5	72.1	376	6 US-08-213-200A-2 Sequence 2, Appl1
14	1539.5	72.1	376	19 US-09-593-316-6 Sequence 6, Appl1
15	1539.5	72.1	376	27 US-09-173-270-2 Sequence 2, Appl1
16	1539.5	72.1	376	28 US-09-994-427A-6 Sequence 6, Appl1
17	1539.5	72.1	376	28 US-09-995-419A-9 Sequence 9, Appl1
18	1532	71.7	363	16 US-09-230-091-4 Sequence 4, Appl1
19	1523.5	71.3	376	28 US-09-994-427A-13 Sequence 13, Appl1
20	1513	70.8	375	16 US-09-230-091-2 Sequence 2, Appl1
21	1510	70.7	371	13 US-08-984-900-10 Sequence 10, Appl1
22	1510	70.7	371	19 US-09-593-316-8 Sequence 8, Appl1
23	1510	70.7	371	28 US-09-994-427A-10 Sequence 10, Appl1
24	1493	69.9	371	5 US-08-188-607B-10 Sequence 10, Appl1
25	1493	69.9	371	7 US-08-378-617-10 Sequence 10, Appl1
26	1476.5	69.1	368	5 US-08-188-607B-11 Sequence 11, Appl1
27	1476.5	69.1	368	7 US-08-378-617-11 Sequence 11, Appl1
28	1476.5	69.1	368	13 US-08-984-900-11 Sequence 11, Appl1
29	1476.5	69.1	368	19 US-09-593-316-4 Sequence 4, Appl1
30	1476.5	69.1	368	28 US-09-994-427A-9 Sequence 9, Appl1
31	1471	68.9	369	19 US-09-593-316-2 Sequence 2, Appl1
32	1471	68.9	369	28 US-09-994-427A-8 Sequence 8, Appl1
33	1471	68.9	369	28 US-09-995-419A-7 Sequence 7, Appl1
34	1431.5	67.0	354	16 US-09-230-091-6 Sequence 6, Appl1
35	1428.5	66.9	342	16 US-09-230-091-8 Sequence 8, Appl1
36	945	44.2	227	28 US-09-994-427A-7 Sequence 7, Appl1
37	690.5	32.3	335	24 US-60-178-574-9 Sequence 9, Appl1
38	688	32.2	353	28 US-09-994-427A-3 Sequence 3, Appl1
39	688	32.2	353	28 US-09-995-419A-11 Sequence 11, Appl1
40	687.5	32.2	354	28 US-09-994-427A-5 Sequence 5, Appl1
41	687.5	32.2	354	28 US-09-995-419A-13 Sequence 13, Appl1
42	673	31.5	347	18 US-09-488-725A-2396 Sequence 2396, Ap
43	646	30.2	402	18 US-09-488-725A-5968 Sequence 5968, Ap
44	632	29.6	197	19 US-09-593-316-12 Sequence 12, Appl1
45	482	22.6	211	27 US-09-611-526-3056 Sequence 3056, Ap

ALIGNMENTS

RESULT 1  
US-08-379-040-2  
; Sequence 2, Application US/08379040  
; GENERAL INFORMATION:  
; APPLICANT: Cooper, David K.C.  
; APPLICANT: Koren, Eugen  
; TITLE OF INVENTION: GENETICALLY ENGINEERED ANIMALS FOR USE  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patrea L. Pabst  
; STREET: 1100 Peachtree Street, Suite 2800  
; CITY: Atlanta  
; STATE: GA  
; COUNTRY: USA  
; ZIP: 30309-4530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/379,040  
; FILING DATE:

CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/049,817  
FILING DATE: 20-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: BMC100  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404)-815-6508  
TELEFAX: (404)-815-6555  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 394 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORGANISM: Mus musculus  
US-08-379-040-2

Query Match 100.0%; Score 2136; DB 7; Length 394;  
Best Local Similarity 100.0%; Pred. No. 1.7e-212;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSEKIMNVKGVILLMLLIVSTVVVWFVEYV 60  
DB 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSEKIMNVKGVILLMLLIVSTVVVWFVEYV 60  
QY 61 NRIPEVGENRWQKDWPFSPFKNGTHSYQEDNVEGRREKGRNGDRIEEPQLWDWPNKRN 120  
DB 61 NRIPEVGENRWQKDWPFSPFKNGTHSYQEDNVEGRREKGRNGDRIEEPQLWDWPNKRN 120  
QY 121 PDVLTVPWKAPIVWEGTYDTALLEKYATOKLTGVLTVFAVGKYIEHYLEDLFLESADMY 180  
DB 121 PDVLTVPWKAPIVWEGTYDTALLEKYATOKLTGVLTVFAVGKYIEHYLEDLFLESADMY 180  
QY 181 FMVGHVIFYVMIDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
DB 181 FMVGHVIFYVMIDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
QY 241 QHEVDLFCDMDVQVDFQDNFVETLGOLVAQLQAWMYKASPEKTYERRELSAAYIPFGE 300  
DB 241 QHEVDLFCDMDVQVDFQDNFVETLGOLVAQLQAWMYKASPEKTYERRELSAAYIPFGE 300  
QY 301 GDFYHAAIFGGTPTTHLNLTRCFKGLQDKKHIDIEAQWHDHSHLNKYFLFNKPTKILS 360  
DB 301 GDFYHAAIFGGTPTTHLNLTRCFKGLQDKKHIDIEAQWHDHSHLNKYFLFNKPTKILS 360  
QY 361 PEYCDYQIGLPSDIKSVKVAWQTKYNLVRNV 394  
DB 361 PEYCDYQIGLPSDIKSVKVAWQTKYNLVRNV 394

## RESULT 2

US-08-823-489-4  
Sequence 4, Application US/08823489  
GENERAL INFORMATION:  
APPLICANT: LOWE, JOHN B.  
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
P.C.  
STREET: 1755 Jefferson Davis Highway, Fourth Floor  
CITY: Arlington  
STATE: Virginia

COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/823,489  
FILING DATE: 25-MAR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/914,281  
FILING DATE: 20-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Lavalleye, Jean-Paul M. P.  
REGISTRATION NUMBER: 31,451  
REFERENCE/DOCKET NUMBER: 2363-060-55  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)521-4500  
TELEFAX: (703)486-2347  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 394 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-823-489-4

Query Match 100.0%; Score 2136; DB 12; Length 394;  
Best Local Similarity 100.0%; Pred. No. 1.7e-212;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSEKIMNVKGVILLMLLIVSTVVVWFVEYV 60  
DB 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSEKIMNVKGVILLMLLIVSTVVVWFVEYV 60  
QY 61 NRIPEVGENRWQKDWPFSPFKNGTHSYQEDNVEGRREKGRNGDRIEEPQLWDWPNKRN 120  
DB 61 NRIPEVGENRWQKDWPFSPFKNGTHSYQEDNVEGRREKGRNGDRIEEPQLWDWPNKRN 120  
QY 121 PDVLTVPWKAPIVWEGTYDTALLEKYATOKLTGVLTVFAVGKYIEHYLEDLFLESADMY 180  
DB 121 PDVLTVPWKAPIVWEGTYDTALLEKYATOKLTGVLTVFAVGKYIEHYLEDLFLESADMY 180  
QY 181 FMVGHVIFYVMIDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
DB 181 FMVGHVIFYVMIDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
QY 241 QHEVDLFCDMDVQVDFQDNFVETLGOLVAQLQAWMYKASPEKTYERRELSAAYIPFGE 300  
DB 241 QHEVDLFCDMDVQVDFQDNFVETLGOLVAQLQAWMYKASPEKTYERRELSAAYIPFGE 300  
QY 301 GDFYHAAIFGGTPTTHLNLTRCFKGLQDKKHIDIEAQWHDHSHLNKYFLFNKPTKILS 360  
DB 301 GDFYHAAIFGGTPTTHLNLTRCFKGLQDKKHIDIEAQWHDHSHLNKYFLFNKPTKILS 360  
QY 361 PEYCDYQIGLPSDIKSVKVAWQTKYNLVRNV 394  
DB 361 PEYCDYQIGLPSDIKSVKVAWQTKYNLVRNV 394

## RESULT 3

US-09-593-316-10  
Sequence 10, Application US/09593316  
GENERAL INFORMATION:  
APPLICANT: Clark Mr., John  
APPLICANT: Danning, Chris  
TITLE OF INVENTION: Animal Tissue For Xenotransplantation  
FILE REFERENCE: 730/002  
CURRENT APPLICATION NUMBER: US/09/593,316



; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: 60/204,148
; PRIOR FILING DATE: 2000-05-15
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Musca sp.
US-09-593-316-10

Query Match 100.0%; Score 2136; DB 19; Length 394;
Best Local Similarity 100.0%; Pred. No. 1.7e-212;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MITMLQDLHVNKISMSRSKSETSLPSSRSGSQEKIMNVKGVILLMLLIVSTVVVWFVEYV 60
DB 1 MITMLQDLHVNKISMSRSKSETSLPSSRSGSQEKIMNVKGVILLMLLIVSTVVVWFVEYV 60

QY 61 NRIPVEGENRWQKDMWFPSPFKNGTHSYQEDNVGRRKGRNGDRIEEPQLWDFNPKNR 120
DB 61 NRIPVEGENRWQKDMWFPSPFKNGTHSYQEDNVGRRKGRNGDRIEEPQLWDFNPKNR 120

QY 121 PDVLTVPKAPIVWEGTYDTALLEKYATQKLTGLTVFAVGKYIEHYLEDFLESADMY 180
DB 121 PDVLTVPKAPIVWEGTYDTALLEKYATQKLTGLTVFAVGKYIEHYLEDFLESADMY 180

QY 181 FMVGHVIFVYVIMDDTSRMPVHNLPLHLSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240
DB 181 FMVGHVIFVYVIMDDTSRMPVHNLPLHLSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240

QY 241 QHEVDFLCMDVDQVQDNFGVETLGLVAQLQAWYKASPEKTYERRELSAAYIPFGE 300
DB 241 QHEVDFLCMDVDQVQDNFGVETLGLVAQLQAWYKASPEKTYERRELSAAYIPFGE 300

QY 301 GDFYHAAIFGGTPTHLNLTRECFKGILODKKHIDEAQAQWHDHSHLNKYLFLNKPTKILS 360
DB 301 GDFYHAAIFGGTPTHLNLTRECFKGILODKKHIDEAQAQWHDHSHLNKYLFLNKPTKILS 360

QY 361 PEYCWQYQIGLPSDIKSVKAWOTKEYNLVRNV 394
DB 361 PEYCWQYQIGLPSDIKSVKAWOTKEYNLVRNV 394

RESULT 4

US-09-863-475-4
; Sequence 4, Application US/09863475
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
;
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
ADDRESSSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
; COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/863,475
FILING DATE: 24-May-2001
CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/914,281
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Lavallee, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 394 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-863-475-4

Query Match 100.0%; Score 2136; DB 22; Length 394;
Best Local Similarity 100.0%; Pred. No. 1.7e-212;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MITMLQDLHVNKISMSRSKSETSLPSSRSGSQEKIMNVKGVILLMLLIVSTVVVWFVEYV 60
DB 1 MITMLQDLHVNKISMSRSKSETSLPSSRSGSQEKIMNVKGVILLMLLIVSTVVVWFVEYV 60

QY 61 NRIPVEGENRWQKDMWFPSPFKNGTHSYQEDNVGRRKGRNGDRIEEPQLWDFNPKNR 120
DB 61 NRIPVEGENRWQKDMWFPSPFKNGTHSYQEDNVGRRKGRNGDRIEEPQLWDFNPKNR 120

QY 121 PDVLTVPKAPIVWEGTYDTALLEKYATQKLTGLTVFAVGKYIEHYLEDFLESADMY 180
DB 121 PDVLTVPKAPIVWEGTYDTALLEKYATQKLTGLTVFAVGKYIEHYLEDFLESADMY 180

QY 181 FMVGHVIFVYVIMDDTSRMPVHNLPLHLSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240
DB 181 FMVGHVIFVYVIMDDTSRMPVHNLPLHLSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240

QY 241 QHEVDFLCMDVDQVQDNFGVETLGLVAQLQAWYKASPEKTYERRELSAAYIPFGE 300
DB 241 QHEVDFLCMDVDQVQDNFGVETLGLVAQLQAWYKASPEKTYERRELSAAYIPFGE 300

QY 301 GDFYHAAIFGGTPTHLNLTRECFKGILODKKHIDEAQAQWHDHSHLNKYLFLNKPTKILS 360
DB 301 GDFYHAAIFGGTPTHLNLTRECFKGILODKKHIDEAQAQWHDHSHLNKYLFLNKPTKILS 360

QY 361 PEYCWQYQIGLPSDIKSVKAWOTKEYNLVRNV 394
DB 361 PEYCWQYQIGLPSDIKSVKAWOTKEYNLVRNV 394

RESULT 5

US-09-863-475A-4
; Sequence 4, Application US/09863475A
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
;
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
ADDRESSSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
; COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION NUMBER: US/09/863.475A  
FILING DATE: 24-May-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/914,281  
FILING DATE: 20-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Lavalleye, Jean-Paul M. P.  
REGISTRATION NUMBER: 31,451  
REFERENCE/DOCKET NUMBER: 2363-060-55  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)521-4500  
TELEFAX: (703)486-2347  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 394 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-863-475A-4

Query Match 100.0%; Score 2136; DB 27; Length 394;  
Best Local Similarity 100.0%; Pred. No. 1.7e-212; Indels 0; Gaps 0;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MITLQDLHVNKISMSRSKSETSLPSSRSGSQEKIMNVKGVILLMLLIVSTVVVFEYV 60  
Db 1 MITLQDLHVNKISMSRSKSETSLPSSRSGSQEKIMNVKGVILLMLLIVSTVVVFEYV 60  
Qy 61 NRIPEVGENRQKQDWPFQNFVETLGQVLAQLQAWYKASPEKFTYERRELSAAYIPGE 120  
Db 61 NRIPEVGENRQKQDWPFQNFVETLGQVLAQLQAWYKASPEKFTYERRELSAAYIPGE 120  
Qy 121 PDVLTVPWKAPIVWEGTYDTALLEKYATOKLTGTVFAVGKYIEHYLEDPLESADMY 180  
Db 121 PDVLTVPWKAPIVWEGTYDTALLEKYATOKLTGTVFAVGKYIEHYLEDPLESADMY 180  
Qy 181 FMVGHVIFYVMIDTTSRMPVHLNPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
Db 181 FMVGHVIFYVMIDTTSRMPVHLNPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
Qy 241 QHEVDLFECMDVDQVFNQFVETLGQVLAQLQAWYKASPEKFTYERRELSAAYIPGE 300  
Db 241 QHEVDLFECMDVDQVFNQFVETLGQVLAQLQAWYKASPEKFTYERRELSAAYIPGE 300  
Qy 301 GDFYHAAIFGCTTPTTHILNLTRECFCGTLQDKKHDIEAQWHDHSHLNKYFLFNKPTK 360  
Db 301 GDFYHAAIFGCTTPTTHILNLTRECFCGTLQDKKHDIEAQWHDHSHLNKYFLFNKPTK 360  
Qy 361 PEYCDWQIGLPSDIKSVKVAQWTKKEYNLVRNV 394  
Db 361 PEYCDWQIGLPSDIKSVKVAQWTKKEYNLVRNV 394

RESULT 6  
US-09-946-034-2  
Sequence 2, Application US/09946034  
GENERAL INFORMATION:  
APPLICANT: Cooper, David K.C.  
Koren, Eugen  
TITLE OF INVENTION: GENETICALLY ENGINEERED ANIMALS FOR USE  
AS ORGAN DONORS  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patrea L. Pabst  
STREET: 1100 Peachtree Street, Suite 2800

CITY: Atlanta  
STATE: GA  
COUNTRY: USA  
ZIP: 30309-4530  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/946.034  
FILING DATE: 04-Sep-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/049,817  
FILING DATE: 20-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: BMC100  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404)-815-6508  
TELEFAX: (404)-815-6555  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 394 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Mus musculus  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-946-034-2

Query Match 100.0%; Score 2136; DB 27; Length 394;  
Best Local Similarity 100.0%; Pred. No. 1.7e-212; Indels 0; Gaps 0;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MITLQDLHVNKISMSRSKSETSLPSSRSGSQEKIMNVKGVILLMLLIVSTVVVFEYV 60  
Db 1 MITLQDLHVNKISMSRSKSETSLPSSRSGSQEKIMNVKGVILLMLLIVSTVVVFEYV 60  
Qy 61 NRIPEVGENRQKQDWPFQNFVETLGQVLAQLQAWYKASPEKFTYERRELSAAYIPGE 120  
Db 61 NRIPEVGENRQKQDWPFQNFVETLGQVLAQLQAWYKASPEKFTYERRELSAAYIPGE 120  
Qy 121 PDVLTVPWKAPIVWEGTYDTALLEKYATOKLTGTVFAVGKYIEHYLEDPLESADMY 180  
Db 121 PDVLTVPWKAPIVWEGTYDTALLEKYATOKLTGTVFAVGKYIEHYLEDPLESADMY 180  
Qy 181 FMVGHVIFYVMIDTTSRMPVHLNPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
Db 181 FMVGHVIFYVMIDTTSRMPVHLNPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
Qy 241 QHEVDLFECMDVDQVFNQFVETLGQVLAQLQAWYKASPEKFTYERRELSAAYIPGE 300  
Db 241 QHEVDLFECMDVDQVFNQFVETLGQVLAQLQAWYKASPEKFTYERRELSAAYIPGE 300  
Qy 301 GDFYHAAIFGCTTPTTHILNLTRECFCGTLQDKKHDIEAQWHDHSHLNKYFLFNKPTK 360  
Db 301 GDFYHAAIFGCTTPTTHILNLTRECFCGTLQDKKHDIEAQWHDHSHLNKYFLFNKPTK 360  
Qy 361 PEYCDWQIGLPSDIKSVKVAQWTKKEYNLVRNV 394  
Db 361 PEYCDWQIGLPSDIKSVKVAQWTKKEYNLVRNV 394

RESULT 7  
US-09-994-427A-11  
Sequence 11, Application US/09994427A

```

; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Schiff, J. Michael
; TITLE OF INVENTION: GLYCOSYLTRANSFERASE VECTORS FOR TREATING CANCER
; FILE REFERENCE: 083.002
; CURRENT APPLICATION NUMBER: US/09/994,427A
; CURRENT FILING DATE: 2002-02-26
; PRIOR APPLICATION NUMBER: 60/253,395
; PRIOR FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-994-427A-11

Query Match          92.2%; Score 1969; DB 28; Length 359;
Best Local Similarity 100.0%; Pred. No. 3.5e-195;
Matches 359; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 36 MNVKGKILLMLIVSTVVVFWFVWVRIPEVGENRWQKQKQHWFPKNGTHSYQEDNVEG 95
Db 1 MNVKGKILLMLIVSTVVVFWFVWVRIPEVGENRWQKQKQHWFPKNGTHSYQEDNVEG 60
Qy 96 RREKGRNGRIEPPQLMDWPNKPNRPDLVTPWKAPIVWEGTYDTALLEKYATOKLTV 155
Db 61 RREKGRNGRIEPPQLMDWPNKPNRPDLVTPWKAPIVWEGTYDTALLEKYATOKLTV 120
Qy 156 GLTVFAVGKYEIHYLEDFLESADMYFMVGHVIFVYVMIDTTSRMPVHLNPLHSLQVFEI 215
Db 121 GLTVFAVGKYEIHYLEDFLESADMYFMVGHVIFVYVMIDTTSRMPVHLNPLHSLQVFEI 180
Qy 216 RSEKRWODISMWRMKTIGEHLAHIQHEVDFLECMVDVQVDFQDFGVTETLGQLVAQLQAW 275
Db 181 RSEKRWODISMWRMKTIGEHLAHIQHEVDFLECMVDVQVDFQDFGVTETLGQLVAQLQAW 240
Qy 276 WYKASPEKFTYERRELSAAYIPFEGDFYVYHAAIFGGTPTTHILNLTRECFKGLQKKHD 335
Db 241 WYKASPEKFTYERRELSAAYIPFEGDFYVYHAAIFGGTPTTHILNLTRECFKGLQKKHD 300
Qy 336 IEAQWHDHSHLNKYFLFNKPTKILSPYCWYDQIGLPSDIKSVKVAWQTKYENLVRNV 394
Db 301 IEAQWHDHSHLNKYFLFNKPTKILSPYCWYDQIGLPSDIKSVKVAWQTKYENLVRNV 359

RESULT      8
US-08-188-607B-12
; Sequence 12, Application US/08188607B
; GENERAL INFORMATION:
; APPLICANT: d'Apice, Anthony J.F.
; APPLICANT: Pearce, Martin J.
; APPLICANT: Robins, Allan J.
; APPLICANT: Crawford, Robert J.
; APPLICANT: Rathjen, Peter D.
; TITLE OF INVENTION: MATERIALS AND METHODS FOR MANAGEMENT OF HYPERACUTE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patterson & Keough, P.A.
; STREET: 527 Marquette Avenue South
; CITY: Minneapolis
; STATE: Minnesota
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh System 7.0
; SOFTWARE: Macintosh Text File
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/188,607B
; FILING DATE: 27-01-1994
```

```

; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark S. Ellinger
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 1175.04-US-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612/349-5740
; TELEFAX: 612/349-9266
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 371
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: protein
; HYPOTHEICAL: no
; ANTI-SENSE:
; FRAGMENT TYPE: Entire protein
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
; STRAIN:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; FEATURE:
; NAME/KEY: Murine alpha-1,3-Galactosyltransferase amino acid sequence dedu
; PUBLICATION INFORMATION:
; AUTHORS: Joiasse, D. H.
; AUTHORS: Shaper, N.L.
; AUTHORS: Kim, D.
; AUTHORS: van den Eijnden, D.H.
; AUTHORS: Shaper, J. H.
; TITLE: Murine alpha-1,3-galactosyltransferase: A single gene locus specifi
; JOURNAL: The Journal of Biological Chemistry
; VOLUME: 267
; ISSUE: 8
; PAGES: 5534-5541
; DATE: 1992
US-08-188-607B-12

Query Match          91.3%; Score 1950; DB 5; Length 371;
Best Local Similarity 96.5%; Pred. No. 3.4e-193;
Matches 358; Conservative 1; Mismatches 0; Indels 12; Gaps 1;

Qy 36 MNVKGKILLMLIVSTVVVFWFVWVRIPEVGENRWQKQKQHWFPKNGTHSYQEDNVEG 83
Db 1 MNVKGKILLMLIVSTVVVFWFVWVRIPEVGENRWQKQKQHWFPKNGTHSYQEDNVEG 60
Qy 84 GTHSYQEDNVEGRKGRNGRIEPPQLMDWPNKPNRPDLVTPWKAPIVWEGTYDTAL 143
Db 61 GTHSYQEDNVEGRKGRNGRIEPPQLMDWPNKPNRPDLVTPWKAPIVWEGTYDTAL 120
Qy 144 LEKYATQKLTVGLTVFAVGKYEIHYLEDFLESADMYFMVGHVIFVYVMIDTTSRMPVWH 203
Db 121 LEKYATQKLTVGLTVFAVGKYEIHYLEDFLESADMYFMVGHVIFVYVMIDTTSRMPVWH 180
Qy 204 LNPPLHSLQVFEIRSEKRWQDISMRMKTIGEHLAHIQHEVDFLECMVDVQVDFQDFGVE 263
Db 181 LNPPLHSLQVFEIRSEKRWQDISMRMKTIGEHLAHIQHEVDFLECMVDVQVDFQDFGVE 240
Qy 264 TLGOLVAQLQAWYKASPEKFTYERRELSAAYIPFEGDFYVYHAAIFGGTPTTHILNLTRE 323
Db 241 TLGOLVAQLQAWYKASPEKFTYERRELSAAYIPFEGDFYVYHAAIFGGTPTTHILNLTRE 300
Qy 324 CFKGILODKKHDIQAQWHDHSHLNKYFLFNKPTKILSPYCWYDQIGLPSDIKSVKVAWQ 383
```

Db 301 CFKGLQDKKHIDIEAOWHDESHLNKYLFLFNKPTKILSPYCWQYQIGLPSDIKSVKVAWQ 360  
Qy 384 TKEYNLVRNV 394  
Db 361 TKEYNLVRNV 371  
RESULT 9  
US-08-617-12  
; Sequence 12, Application US/08378617  
; GENERAL INFORMATION:  
; APPLICANT: d'Apice, Anthony J.F.  
; APPLICANT: Pearse, Martin J.  
; APPLICANT: Robins, Allan J.  
; APPLICANT: Crawford, Robert J.  
; APPLICANT: Rathjen, Peter D.  
; TITLE OF INVENTION: MATERIALS AND METHODS FOR MANAGEMENT OF  
; TITLE OF INVENTION: HYPERACUTE REJECTION IN HUMAN XENOTRANSPLANTATION  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 120 South Sixth Street, Suite 2500  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/378.617  
; FILING DATE: 26-JAN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ellinger, Mark S.  
; REGISTRATION NUMBER: 34,812  
; REFERENCE/DOCKET NUMBER: 06868/005001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (612) 335-5070  
; TELEFAX: (612) 288-9696  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 371 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-617-12

Query Match 91.3%; Score 1950; DB 7; Length 371;  
Best Local Similarity 96.5%; Pred. No. 3.4e-193;  
Matches 358; Conservative 1; Mismatches 0; Indels 12; Gaps 1;  
Qy 36 MNVKGKIVLLMLIVSTVVVWFVEYVN-----RIPEVGENRQKQDMWFPSPFN 83  
Db 1 MNVKGKIVLLMLIVSTVVVWFVEYVNSPDGSLWYHTKIPEVGENRQKQDMWFPSPFN 60  
Qy 84 GTHSYQEDNVGREGKGRNGRIEPEQLWDWPNKRNPDVLTVPWKAPIVWEGTYDTAL 143  
Db 61 GTHSYQEDNVGREGKGRNGRIEPEQLWDWPNKRNPDVLTVPWKAPIVWEGTYDTAL 120  
Qy 144 LEKYATQKLVGLTVFAVGKYIEHYLEDLFLESADMFVWGHRRVIFYVMIIDTSMRPVH 203  
Db 121 LEKYATQKLVGLTVFAVGKYIEHYLEDLFLESADMFVWGHRRVIFYVMIIDTSMRPVH 180  
Qy 204 LNPVLSQVFEIRSEKRWQDISMMRMKTIGRHLAHIQHEVDFLFCMDVDQVQDNFEGVE 263  
Db 181 LNPVLSQVFEIRSEKRWQDISMMRMKTIGRHLAHIQHEVDFLFCMDVDQVQDNFEGVE 240  
Qy 264 TLGQLVAQLQAWMYKASPEKTYERRELSAAYIPFEGGDFYHAAIFGGTPTHILNLTRE 323

Db 241 TLGQLVAQLQAWMYKASPEKTYERRELSAAYIPFEGGDFYHAAIFGGTPTHILNLTRE 300  
Qy 324 CFKGLQDKKHIDIEAOWHDESHLNKYLFLFNKPTKILSPYCWQYQIGLPSDIKSVKVAWQ 383  
Db 301 CFKGLQDKKHIDIEAOWHDESHLNKYLFLFNKPTKILSPYCWQYQIGLPSDIKSVKVAWQ 360  
Qy 384 TKEYNLVRNV 394  
Db 361 TKEYNLVRNV 371  
RESULT 10  
US-08-984-900-12  
; Sequence 12, Application US/08984900  
; GENERAL INFORMATION:  
; APPLICANT: d'Apice, Anthony J.F.  
; APPLICANT: Pearse, Martin J.  
; APPLICANT: Robins, Allan J.  
; APPLICANT: Crawford, Robert J.  
; APPLICANT: Rathjen, Peter D.  
; TITLE OF INVENTION: MATERIALS AND METHODS FOR MANAGEMENT OF  
; TITLE OF INVENTION: HYPERACUTE REJECTION IN HUMAN XENOTRANSPLANTATION  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 120 South Sixth Street, Suite 2500  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/984,900  
; FILING DATE:  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/378,617  
; FILING DATE: 26-JAN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ellinger, Mark S.  
; REGISTRATION NUMBER: 34,812  
; REFERENCE/DOCKET NUMBER: 06868/005001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (612) 335-5070  
; TELEFAX: (612) 288-9696  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 371 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-984-900-12

Query Match 91.3%; Score 1950; DB 13; Length 371;  
Best Local Similarity 96.5%; Pred. No. 3.4e-193;  
Matches 358; Conservative 1; Mismatches 0; Indels 12; Gaps 1;  
Qy 36 MNVKGKIVLLMLIVSTVVVWFVEYVN-----RIPEVGENRQKQDMWFPSPFN 83  
Db 1 MNVKGKIVLLMLIVSTVVVWFVEYVNSPDGSLWYHTKIPEVGENRQKQDMWFPSPFN 60  
Qy 84 GTHSYQEDNVGREGKGRNGRIEPEQLWDWPNKRNPDVLTVPWKAPIVWEGTYDTAL 143  
Db 61 GTHSYQEDNVGREGKGRNGRIEPEQLWDWPNKRNPDVLTVPWKAPIVWEGTYDTAL 120  
Qy 144 LEKYATQKLVGLTVFAVGKYIEHYLEDLFLESADMFVWGHRRVIFYVMIIDTSMRPVH 203

Db 121 LEKYATOKLTGVTGAVGKYIEHYLEDFLESADMYFMVGHVRIFYVMIDDTSRMPVVH 180  
Qy 204 LNPILSLQVFEIRSEKRWQDISMMRMKTIGEHILAHQHEVDVFLFCMDVDQVFOQDFGVE 263  
Db 181 LNPILSLQVFEIRSEKRWQDISMMRMKTIGEHILAHQHEVDVFLFCMDVDQVFOQDFGVE 240  
Qy 264 TLGOLVAQLQAWMYKASPEKFTYERRELSAAYIPFGGDFYYHAAIFGGTPTTHILNLTRE 323  
Db 241 TLGOLVAQLQAWMYKASPEKFTYERRELSAAYIPFGGDFYYHAAIFGGTPTTHILNLTRE 300  
Qy 324 CFKGILODKKHDIQAQWHDHSHLNKYLFLNKPTKILSPCYCWDYQIGLPSDIKSVKVAWO 383  
Db 301 CFKGILODKKHDIQAQWHDHSHLNKYLFLNKPTKILSPCYCWDYQIGLPSDIKSVKVAWO 360  
Qy 384 TKEYNLVRNNV 394  
Db 361 TKEYNLVRNNV 371

## RESULT 11

US-09-870-759-74  
; Sequence 74, Application US/09870759  
; GENERAL INFORMATION:  
; APPLICANT: TERMAN, David S  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE  
; FILE REFERENCE: 870759  
; CURRENT APPLICATION NUMBER: US/09/870,759  
; CURRENT FILING DATE: 2002-01-14  
; PRIOR APPLICATION NUMBER: US 60/208,128  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 166  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 74  
; LENGTH: 371  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-870-759-74

Query Match 91.3%; Score 1950; DB 28; Length 371;  
Best Local Similarity 96.5%; Pred. No. 3.4e-193;  
Matches 358; Conservative 1; Mismatches 0; Indels 12; Gaps 1;

Qy 36 MNVKGKVVLLMLIVSTVVVFWVEYVN-----RIPEVGENRWQDMWFPSPFKN 83  
Db 1 MNVKGKVVLLMLIVSTVVVFWVEYVNSPDGSLWYHTKIPVGENRWQDMWFPSPFKN 60  
Qy 84 GTHSYQEDNVEGRKGRNGDRIEPPQLWDFNPKNRPDLVTPWKPAPVWEGTYDTAL 143  
Db 61 GTHSYQEDNVEGRKGRNGDRIEPPQLWDFNPKNRPDLVTPWKPAPVWEGTYDTAL 120  
Qy 144 LEKYATOKLTGVTGAVGKYIEHYLEDFLESADMYFMVGHVRIFYVMIDDTSRMPVVH 203  
Db 121 LEKYATOKLTGVTGAVGKYIEHYLEDFLESADMYFMVGHVRIFYVMIDDTSRMPVVH 180  
Qy 204 LNPILSLQVFEIRSEKRWQDISMMRMKTIGEHILAHQHEVDVFLFCMDVDQVFOQDFGVE 263  
Db 181 LNPILSLQVFEIRSEKRWQDISMMRMKTIGEHILAHQHEVDVFLFCMDVDQVFOQDFGVE 240  
Qy 264 TLGOLVAQLQAWMYKASPEKFTYERRELSAAYIPFGGDFYYHAAIFGGTPTTHILNLTRE 323  
Db 241 TLGOLVAQLQAWMYKASPEKFTYERRELSAAYIPFGGDFYYHAAIFGGTPTTHILNLTRE 300  
Qy 324 CFKGILODKKHDIQAQWHDHSHLNKYLFLNKPTKILSPCYCWDYQIGLPSDIKSVKVAWO 383  
Db 301 CFKGILODKKHDIQAQWHDHSHLNKYLFLNKPTKILSPCYCWDYQIGLPSDIKSVKVAWO 360  
Qy 384 TKEYNLVRNNV 394  
Db 361 TKEYNLVRNNV 371

## RESULT 12

US-09-994-427A-12  
; Sequence 12, Application US/09994427A  
; GENERAL INFORMATION:  
; APPLICANT: Geron Corporation  
; APPLICANT: Schiff, J. Michael  
; TITLE OF INVENTION: GLYCOSYLTRANSFERASE VECTORS FOR TREATING CANCER  
; FILE REFERENCE: 083,002  
; CURRENT APPLICATION NUMBER: US/09/994,427A  
; CURRENT FILING DATE: 2002-02-26  
; PRIOR APPLICATION NUMBER: 60/253,395  
; PRIOR FILING DATE: 2000-11-27  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 12  
; LENGTH: 376  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Consensus of mammalian galactosyl transferase sequences - this  
; OTHER INFORMATION: vention  
US-09-994-427A-12

Query Match 72.2%; Score 1542.5; DB 28; Length 376;  
Best Local Similarity 72.9%; Pred. No. 7.9e-151;  
Matches 274; Conservative 42; Mismatches 43; Indels 17; Gaps 2;  
Qy 36 MNVKGKVVLLMLIVSTVVVFWVEYVN-----RIPEVGENRWQDMWFPSPFKN 83  
Db 1 MNVKGKVVLLMLIVSTVVVFWVEYVNSPDGSLWYHTKIPVGENRWQDMWFPSPFKN 60  
Qy 84 GTHSYQEDNVEGRKGRNGDRIEPPQLWDFNPKNRPDLVTPWKPAPVWEGTYDTAL 138  
Db 61 GTHSYQEDNVEGRKGRNGDRIEPPQLWDFNPKNRPDLVTPWKPAPVWEGTYDTAL 120  
Qy 139 YDTALLEKYATOKLTGVTGAVGKYIEHYLEDFLESADMYFMVGHVRIFYVMIDDTSR 198  
Db 121 YNKAILENYYAKOKITVGLTVFAIGRYIEHYLEDFLESADMYFMVGHVRIFYVMIDDTSR 180  
Qy 199 MPVHLNPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAHQHEVDVFLFCMDVDQVFOQ 258  
Db 181 APFIELOPLKSFVFEVPEKRWQDISMMRMKTIGEHILAHQHEVDVFLFCMDVDQVFOQ 240  
Qy 259 NFGVETLGQLVAQLQAWMYKASPEKFTYERRELSAAYIPFGGDFYYHAAIFGGTPTTHIL 318  
Db 241 HFGVETLGQLVAQLQAWMYKADPDFTYERRKESAAVPEGQGFYYHAAIFGGTPTTHIL 300  
Qy 319 NLTRCEKGILODKKHDIQAQWHDHSHLNKYLFLNKPTKILSPCYCWDYQIGLPSDIKSV 378  
Db 301 NITQCEKGILODKKHDIQAQWHDHSHLNKYLFLNKPTKILSPCYCWDYQIGLPSDIKSV 360  
Qy 379 KVAMOTKEYNLVRNNV 394  
Db 361 KLSQTKKEYNLVRNNV 376

## RESULT 13

US-08-213-200A-2  
; Sequence 2, Application US/08213200A  
; GENERAL INFORMATION:  
; APPLICANT: Galili, Dr. Uri  
; APPLICANT: Replik, Dr. Patricia M.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR VACCINES  
; TITLE OF INVENTION: COMPRISING ALPHA-CALACTOSYL EPITOPES  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.  
; STREET: Suite 1800, Two Penn Center Plaza  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19102  
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE:  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/213,200A  
FILING DATE: 15-MAR-1994  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Daniel A. Monaco  
REGISTRATION NUMBER: 30,480  
REFERENCE/DOCKET NUMBER: 8760-2 US  
TELEPHONE: (215) 568-8383  
TELEFAX: (215) 568-5549  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 376 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-213-200A-2

Query Match 72.1%; Score 1539.5; DB 6; Length 376;  
Best Local Similarity 72.6%; Pred. No. 1.6e-150;  
Matches 273; Conservative 43; Mismatches 43; Indels 17; Gaps 2;  
QY 36 MNVKGKILLMLIVSTVVVFWEYVN-----RIPEVGENRWQKDWFPSPFN 83  
DB 1 MNVKGKILLSMLVSTVVVFWEYINSPEGSLWIYHSKNPEVDDSSAQKDWFPSPFN 60  
QY 84 GTHSYQE-----DNVEGRREKGRNGRIEPEQLWDWPNKRNPDVLTVPWKAPVMEGT 138  
DB 61 GIHNYQEEEDTDKEKGRREEQKEDDTTELRLWDWPNKRPVMTVQNKAPVMEGT 120  
QY 139 YDTALLEKYATQKLTVGLTVFAVGKYEIEHYLEDFLESADMYFMVGHVRVIFYVMIDTISR 198  
DB 121 YNKAILENYAKQKITVGLTVFAIGRYIEHYLEEFVTSANRYFMVGHKVFIFYVMVDVSK 180  
QY 199 MPVVHLNPLHSLOVFEIRSEKRWODISMRRMKTIGEHLAHQHEVDVDFLCMDVDVQFOD 258  
DB 181 APTELGLPLRSFKVFEVPEKRWQDISMMRMKTIGEHLAHQHEVDVDFLCMDVDVQFOD 240  
QY 259 NFGVETLQGLVAQLQAWMYKASPEKFTYERRELSAAYIPFEGDFYHYHAAIFGCTPTHIL 318  
DB 241 HFGVETLQGSVAQLQAWMYKADPDFTYERKESAAVIFPGQDFYHYHAAIFGCTPIQVL 300  
QY 319 NLTRCEFGILQDKKHIDIEAQWHDHSHLNKYFLFNKPTKILSPCYWDYQIGLPSDIKSV 378  
DB 301 NITQECFKGILLDKKNDIEAEWHDESHLNKYFLFNKPSKILSPCYWDYHIGLPSDIKT 360  
QY 379 KVAMQTKYNNLVRNNV 394  
DB 361 KLSWQTKYNNLVRNNV 376

RESULT 14  
US-09-593-316-6  
SEQUENCE 6, Application US/09593316  
GENERAL INFORMATION:  
APPLICANT: Clark Mr., John  
TITLE OF INVENTION: Animal Tissue For Xenotransplantation  
FILE REFERENCE: 730/002  
CURRENT APPLICATION NUMBER: US/09/593,316  
PRIOR FILING DATE: 2000-06-13  
PRIOR APPLICATION NUMBER: 60/204,148  
NUMBER OF SEQ ID NOS: 51  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 376

TYPE: PRT  
ORGANISM: Mermoset alpha 1.3-CT  
US-09-593-316-6  
Query Match 72.1%; Score 1539.5; DB 19; Length 376;  
Best Local Similarity 72.6%; Pred. No. 1.6e-150;  
Matches 273; Conservative 43; Mismatches 43; Indels 17; Gaps 2;  
QY 36 MNVKGKILLMLIVSTVVVFWEYVN-----RIPEVGENRWQKDWFPSPFN 83  
DB 1 MNVKGKILLSMLVSTVVVFWEYINSPEGSLWIYHSKNPEVDDSSAQKDWFPSPFN 60  
QY 84 GTHSYQE-----DNVEGRREKGRNGRIEPEQLWDWPNKRNPDVLTVPWKAPVMEGT 138  
DB 61 GIHNYQEEEDTDKEKGRREEQKEDDTTELRLWDWPNKRPVMTVQNKAPVMEGT 120  
QY 139 YDTALLEKYATQKLTVGLTVFAVGKYEIEHYLEDFLESADMYFMVGHVRVIFYVMIDTISR 198  
DB 121 YNKAILENYAKQKITVGLTVFAIGRYIEHYLEEFVTSANRYFMVGHKVFIFYVMVDVSK 180  
QY 199 MPVVHLNPLHSLOVFEIRSEKRWODISMRRMKTIGEHLAHQHEVDVDFLCMDVDVQFOD 258  
DB 181 APTELGLPLRSFKVFEVPEKRWQDISMMRMKTIGEHLAHQHEVDVDFLCMDVDVQFOD 240  
QY 259 NFGVETLQGLVAQLQAWMYKASPEKFTYERRELSAAYIPFEGDFYHYHAAIFGCTPTHIL 318  
DB 241 HFGVETLQGSVAQLQAWMYKADPDFTYERKESAAVIFPGQDFYHYHAAIFGCTPIQVL 300  
QY 319 NLTRCEFGILQDKKHIDIEAQWHDHSHLNKYFLFNKPTKILSPCYWDYQIGLPSDIKSV 378  
DB 301 NITQECFKGILLDKKNDIEAEWHDESHLNKYFLFNKPSKILSPCYWDYHIGLPSDIKT 360  
QY 379 KVAMQTKYNNLVRNNV 394  
DB 361 KLSWQTKYNNLVRNNV 376  
RESULT 15  
US-09-173-270-2  
SEQUENCE 2, Application US/09173270  
GENERAL INFORMATION:  
APPLICANT: GALILI, URI  
TITLE OF INVENTION: REPIK, PATRICIA M.  
COMPOSING ALPHA-GALACTOSYL EPITOPES  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P. C.  
STREET: Suite 1800, Two Penn Center Plaza  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19102  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/173,270  
FILING DATE: 23-Mar-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/704,548  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Monaco, Daniel A.  
REGISTRATION NUMBER: 30,480  
REFERENCE/DOCKET NUMBER: 8760-2 CII  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-8383  
TELEFAX: (215) 568-5549

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-173-270-2

Query Match	72.1%	Score 1539.5;	DB 27;	Length 376;
Best Local Similarity	72.6%	Pred. No. 1.6e-150;		
Matches 273;	Conservative 43;	Mismatches 43;	Indels 17;	Gaps 2;

  

QY	36	MNVKGVILMLIVSTVVVFWFYVN-----RIPEYGENRWQKDMWFPSWFKN	83
DB	1	MNVKGVILSMLVSVIVVFWFYINSPGCSFLVIYHSKNPEYDDSSAQKDMWFGWENN	60
QY	84	GTHSYOE-----DNVEGRREKGRNGRIEPPQLMDWFPNKNRPDLTVTPWKAPIVWEGT	138
DB	61	GIHNYQOEEDTDKEREDEEKEDDTTLRLMDWFPNKKRPEVMTVTQWKAPVWEGT	120
QY	139	YDTALLEKYYATOKLTGLTVFAVGKYIEHYLEDFLESADMYFMVGHRYFYVMIDTTSR	198
DB	121	YNKAILENYAKOKITVGLTVFAIGRYIEHYLEEFVTSANRYFMVGHKIFYVMVDDVSK	180
QY	199	MPVVLNPLHLSLQVFEIRSEKRWQDISMMRMKTIGEHILAHIOHEVDFLCMDVDQVQFD	258
DB	181	APFELGPLRSFKVFEVKPEKRWQDISMMRMKTIGEHILAHIOHEVDFLCMDVDQVQFD	240
QY	259	NFGVETLGQLVAQLQAWMYKASPEKTYERRELSAAYIPFGEGDFYYHAAIFGGTPTHL	318
DB	241	HFGVETLGQSVQALQAWMYKADPDDETYERRKESAAAYIPFGQDFYYHAAIFGGTPTQVL	300
QY	319	NLTRECFCGTLQDKKHIDIEAQWHDSEHLNKYFLFNKPTKILSPCYCWDYQIGLPSDIKS	378
DB	301	NITQECFCGILLDDKNDIEAEWHDESHLNKYFLLNKPSKILSPCYCWDYHIGLPSDIKT	360
QY	379	KVARQTKENYLNVRNV	394
DB	361	KLSMQTKENYLNVRNV	376

Search completed: May 10, 2002, 11:09:25
Job time: 208 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2002, 11:06:22 ; Search time 11.66 Seconds  
(without alignments)  
720.742 Million cell updates/sec

Title: US-09-863-475A-4  
Perfect score: 2136  
Sequence: 1 MITMLQDLHVKNISMSRSK.....IKSVKVAWOTKEYNLVRNV 394

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 92415 seqs, 21329587 residues

Total number of hits satisfying chosen parameters: 92415

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*  
1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	229	10.7	106	6	US-10-115-123-419
2	175	8.2	90	6	US-10-115-123-190
3	175	8.2	90	6	US-10-115-123-240
4	175	8.2	90	6	US-10-115-123-417
5	97	4.5	432	5	US-09-602-777A-140
6	86	4.0	887	1	PCT-US02-09288-20
7	85.5	4.0	368	1	PCT-US02-09944-432
8	85.5	4.0	662	5	US-09-540-209B-8400
9	84	3.9	746	6	US-10-041-007-39
10	84	3.9	795	6	US-10-041-007-37
11	84	3.9	814	6	US-10-041-007-35
12	84	3.9	873	6	US-10-041-007-2
13	84	3.9	873	6	US-10-041-007-33
14	84	3.9	873	6	US-10-041-018-398
15	82.5	3.9	696	5	US-09-540-209B-9947
16	79.5	3.7	379	5	US-09-540-209B-6269
17	78.5	3.7	417	5	US-09-540-209B-8036
18	77.5	3.6	2214	1	PCT-US02-07826-300
19	77.5	3.6	2214	1	PCT-US02-09671-810
20	77.5	3.6	2214	1	PCT-US02-09671-811
21	77.5	3.6	2214	1	PCT-US02-09671-812
22	77.5	3.6	2214	1	PCT-US02-09671-813
23	77.5	3.6	2214	1	PCT-US02-09671-814
24	77.5	3.6	2214	6	US-10-097-340-300
25	76	3.6	249	5	US-09-540-209B-5225
26	75.5	3.5	690	5	US-09-540-209B-6647

27	75.5	3.5	1722	1	PCT-US02-09671-666	Sequence 666, App
28	75.5	3.5	1722	1	PCT-US02-09671-667	Sequence 667, App
29	75.5	3.5	1722	1	PCT-US02-09671-670	Sequence 670, App
30	75.5	3.5	1722	1	PCT-US02-09671-671	Sequence 671, App
31	75.5	3.5	1722	1	PCT-US02-09671-672	Sequence 672, App
32	75.5	3.5	1722	1	PCT-US02-09671-673	Sequence 673, App
33	75.5	3.5	1743	1	PCT-US02-09671-668	Sequence 668, App
34	75	3.5	224	5	US-09-540-209B-6438	Sequence 6438, App
35	75	3.5	471	5	US-09-573-655B-945	Sequence 945, App
36	75	3.5	1428	5	US-09-573-655B-1944	Sequence 1944, App
37	74.5	3.5	392	5	US-09-540-209B-9064	Sequence 9064, App
38	74.5	3.5	406	5	US-09-591-521-6	Sequence 6, Appl
39	74.5	3.5	837	5	US-09-540-209B-8217	Sequence 8217, App
40	74.5	3.5	1307	5	US-09-573-655B-1888	Sequence 1888, App
41	74	3.5	468	5	US-09-540-209B-8955	Sequence 8955, App
42	74	3.5	469	1	PCT-US02-10824-164	Sequence 164, App
43	74	3.5	469	1	PCT-US02-10421-2912	Sequence 2912, App
44	74	3.5	469	6	US-10-112-699-2912	Sequence 2912, App
45	74	3.5	485	5	US-09-540-209B-5981	Sequence 5981, App

ALIGNMENTS

RESULT 1

US-10-115-123-419  
; Sequence 419, Application US/10115123  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 94 Human Secreted Proteins  
; FILE REFERENCE: P2029G30AP1D2  
; CURRENT APPLICATION NUMBER: US/10115,123  
; PRIOR FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: PCT/US99/13418  
; PRIOR FILING DATE: 1999-06-15  
; PRIOR APPLICATION NUMBER: 60/089,507  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,508  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,509  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,510  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/090,112  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090,113  
; PRIOR FILING DATE: 1998-06-22  
; NUMBER OF SEQ ID NOS: 532  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 419  
; LENGTH: 106  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-115-123-419

Query Match 10.7%; Score 229; DB 6; Length 106;  
Best Local Similarity 52.4%; Pred. No. 1.3e-14;  
Matches 43; Conservative 12; Mismatches 15; Indels 12; Gaps 1;

Qy 32 OEKTMNVKGVKILLMLIVSTVVVWFVEYVN-----RIPEVGENRWQKDMFPPS 79

Db 3 QEKIRNVKGVKILSMVSVTVIIVFWEFINSTEGSFLWIVHSKNPEVDSSAQKGMWFLS 62

Qy 80 WFKNGTHSYQEDNVGEGRKGR 101

Db 63 WFNNGIHNYQOGEEDIDKEKGR 84

RESULT 2

US-10-115-123-190  
; Sequence 190, Application US/10115123  
; GENERAL INFORMATION:

APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 94 Human Secreted Proteins  
; FILE REFERENCE: P2029G30AP1D2  
; CURRENT APPLICATION NUMBER: US/10/115.123  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: PCT/US99/13418  
; PRIOR FILING DATE: 1999-06-15  
; PRIOR APPLICATION NUMBER: 60/089,507  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,508  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,509  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,510  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/090,112  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090,113  
; PRIOR FILING DATE: 1998-06-22  
; NUMBER OF SEQ ID NOS: 532  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 190  
; LENGTH: 90  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-115-123-190

Query Match 8.2% Score 175; DB 6; Length 90;  
Best Local Similarity 45.6% Pred. No. 1.2e-09;  
Matches 31; Conservative 12; Mismatches 13; Indels 12; Gaps 1;  
  
Qy 46 MLIVSTVVVFEWYVN-----RIPEVGENRQKDWFPSPFKNGTHSYQEDNV 93  
||:||||:||||:| : ||| : || ||| ||| ||| ||| :  
Db 1 MLVSTVIIVFEWFINSTEGSFLWIYHSKNPEVDDSSAQKGMFLSWFNNGIHNYYQGE 60  
: ||| : || ||| ||| ||| ||| :  
  
Qy 94 EGRREKGR 101  
: |||||  
Db 61 DIDKEKGR 68

RESULT 3  
US-10-115-123-240  
; Sequence 240, Application US/10115123  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 94 Human Secreted Proteins  
; FILE REFERENCE: P2029G30AP1D2  
; CURRENT APPLICATION NUMBER: US/10/115.123  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: PCT/US99/13418  
; PRIOR FILING DATE: 1999-06-15  
; PRIOR APPLICATION NUMBER: 60/089,507  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,508  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,509  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,510  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/090,112  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090,113  
; PRIOR FILING DATE: 1998-06-22  
; NUMBER OF SEQ ID NOS: 532  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 240  
; LENGTH: 90  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-115-123-240

Query Match 8.2% Score 175; DB 6; Length 90;  
Best Local Similarity 45.6% Pred. No. 1.2e-09;  
Matches 31; Conservative 12; Mismatches 13; Indels 12; Gaps 1;  
  
Qy 46 MLIVSTVVVFEWYVN-----RIPEVGENRQKDWFPSPFKNGTHSYQEDNV 93  
||:||||:||||:| : ||| : || ||| ||| ||| ||| :  
Db 1 MLVSTVIIVFEWFINSTEGSFLWIYHSKNPEVDDSSAQKGMFLSWFNNGIHNYYQGE 60  
: ||| : || ||| ||| ||| ||| :  
  
Qy 94 EGRREKGR 101  
: |||||  
Db 61 DIDKEKGR 68  
  
RESULT 4  
US-10-115-123-417  
; Sequence 417, Application US/10115123  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 94 Human Secreted Proteins  
; FILE REFERENCE: P2029G30AP1D2  
; CURRENT APPLICATION NUMBER: US/10/115.123  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: PCT/US99/13418  
; PRIOR FILING DATE: 1999-06-15  
; PRIOR APPLICATION NUMBER: 60/089,507  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,508  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,509  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,510  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/090,112  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090,113  
; PRIOR FILING DATE: 1998-06-22  
; NUMBER OF SEQ ID NOS: 532  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 417  
; LENGTH: 90  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-115-123-417

Query Match 8.2% Score 175; DB 6; Length 90;  
Best Local Similarity 45.6% Pred. No. 1.2e-09;  
Matches 31; Conservative 12; Mismatches 13; Indels 12; Gaps 1;  
  
Qy 46 MLIVSTVVVFEWYVN-----RIPEVGENRQKDWFPSPFKNGTHSYQEDNV 93  
||:||||:||||:| : ||| : || ||| ||| ||| ||| :  
Db 1 MLVSTVIIVFEWFINSTEGSFLWIYHSKNPEVDDSSAQKGMFLSWFNNGIHNYYQGE 60  
: ||| : || ||| ||| ||| ||| :  
  
Qy 94 EGRREKGR 101  
: |||||  
Db 61 DIDKEKGR 68  
  
RESULT 5  
US-09-602-777A-140  
; Sequence 140, Application US/09602777A  
; GENERAL INFORMATION:  
; APPLICANT: Pompejus, Markus  
; APPLICANT: Kroger, Burkhard  
; APPLICANT: Schroder, Hartwig  
; APPLICANT: Zelder, Oskar  
; APPLICANT: Haberer, Gregor  
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS  
; FILE REFERENCE: BGI-128CP  
; CURRENT APPLICATION NUMBER: US/09/602,777A  
; CURRENT FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: US 60/141031

PRIOR FILING DATE: 1999-06-25  
PRIOR APPLICATION NUMBER: DE 19931636.8  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19932125.6  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932126.4  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932127.2  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932128.0  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932129.9  
PRIOR FILING DATE: 1999-07-19  
PRIOR APPLICATION NUMBER: DE 19932226.0  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932920.6  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19932922.2  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19932924.9  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19932928.1  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19932930.3  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19932933.8  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19932935.4  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19932973.7  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19933002.6  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19933003.4  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19933005.0  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19933006.9  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19941378.9  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19941379.7  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19941390.8  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19941391.6  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19942088.2  
PRIOR FILING DATE: 1999-09-03  
NUMBER OF SEQ ID NOS: 442  
SEQ ID NO 140  
LENGTH: 432  
TYPE: PRT  
ORGANISM: Corynebacterium glutamicum  
US-09-602-777A-140

Query Match 4.5% Score 97; DB 5; Length 432;  
Best Local Similarity 18.6% Pred. No. 0.19;  
Matches 27; Conservative 27; Mismatches 53; Indels 38; Gaps 3;  
Qy 12 KISMRSKSETSLPSSRSGSGOEKTMNVKGVILLMLIVS-----TTVV 53  
Db 209 KILQAEGEKHASILNABEAERQAMTLRAEGEARAQLQAEARAIQVNAAIKSAKLTP 268  
Qy 54 VVFEYVNRIPVEGKRWKNDWFPSPFNK-----GTHSQEDNVGEGR 98  
Db 269 VLAYOYLEKLPKIAEAGNASKMWVTPSQSDSLEGFAGKQFAGKADGAEVFRPEPNTVDETR 328  
Qy 99 KGRNGDRIEPQLDWDFNPKNRDPV 123  
Db 329 DIANADNVE-----DWFSTESDPEI 348

RESULT 6

PCT-US02-09288-20  
Sequence 20, Application PC/TUS0209288  
GENERAL INFORMATION:

APPLICANT: INCYTE GENOMICS, INC.  
APPLICANT: HAFALIA, April J.A.

APPLICANT: TANG, Y. Tom  
APPLICANT: YUE, Henry

APPLICANT: KHAN, Farrah A.  
APPLICANT: ISON, Craig H.

APPLICANT: BAUGHN, Mariah R.  
APPLICANT: WARREN, Bridget A.

APPLICANT: DUGGAN, Brendan M.  
APPLICANT: THANGAVELU, Kavitha

APPLICANT: HONCHELL, Cynthia D.  
APPLICANT: AZIMZAI, Yalda

APPLICANT: ELLIOTT, Vicki S.  
APPLICANT: BURFORD, Neil

APPLICANT: DING, Li  
APPLICANT: YUE, Huibin

APPLICANT: BECHA, Shanya  
APPLICANT: EMERLING, Brooke M.

APPLICANT: RICHARDSON, Thomas W.  
APPLICANT: LEE, Soo Yeun

APPLICANT: BANDMAN, Olga  
APPLICANT: LAL, Preeti G.

APPLICANT: LEE, Sally  
APPLICANT: GIETZEN, Kimberly J.

APPLICANT: WALIA, Narinder K.  
APPLICANT: GRIFFIN, Jennifer A.

APPLICANT: LEE, Ernestine A.  
APPLICANT: SWARNAKAR, Anita

APPLICANT: RING, Huijun Z.  
APPLICANT: JONES, Karen Anne

TITLE OF INVENTION: CYTOSKELETON-ASSOCIATED PROTEINS  
FILE REFERENCE: PF-0918 PCT

CURRENT APPLICATION NUMBER: PCT/US02/09288  
CURRENT FILING DATE: 2002-03-25

PRIOR APPLICATION NUMBER: 60/280,508; 60/281,323; 60/283,769; 60/288,609; 60/290,5  
60/291,870; 60/294,451

PRIOR FILING DATE: 2001-03-29; 2001-04-03; 2001-04-13; 2001-05-04; 2001-05-10;  
2001-05-18; 2001-05-29

NUMBER OF SEQ ID NOS: 56  
SOFTWARE: PERL Program

SEQ ID NO 20  
LENGTH: 867

TYPE: PRT  
ORGANISM: Homo sapiens

FEATURE:  
NAME/KEY: misc\_feature

OTHER INFORMATION: Incyte ID No: 1990526CD1  
PCT-US02-09288-20

Query Match 4.0% Score 86; DB 1; Length 887;  
Best Local Similarity 19.0% Pred. No. 5.3;  
Matches 82; Conservative 58; Mismatches 138; Indels 154; Gaps 21;

Qy 32 QEKTMNVKGVILLMLIVSTVVVFEYVNRIPVEGKRWKQD----- 74  
Db 83 QEKISELQN-----YQRINHFPGMGE-ICKKDFLARNMTMKIKSRPLDY 125  
Qy 75 -----WVFESWFKNGTHSQEDNVGEGRKQ-----RNGDR----- 105  
Db 126 TFVPRTWIFPAEYTO-QFQYVKELKKRQKQTFIVKPAANGAMGHCISLRNGDKLPQSDH 184  
Qy 106 -----IEEPQLDWDFNPKNRDPVLTVT--PWKAPIVMEGYTDYFALLEKYATO-----K 152  
Db 185 LIVQEIYEKPFMEGYKFDLRIYILVTSQDPLKIFLYHDGLVRMG-TEKYIPPNESNLTO 243  
Qy 153 LTVGLTVFVANGKYIEHYLEDFLESADMYFVWGHVIFVYVIMDDTSRPPVHILNPLHSLOV 212

```
Db 244 LYMHLTNTSVKHNHFRDRTENKG-----SKRSIKWTFEP--LQA 283
Qy 213 FEIRSEKRWODISMRRMKTIGEHLA--HIQHEVDVFLFCMDVDQVQDNFGVETLG---Q 267
Db 284 NOHDVAKFMSDISLVKTL--IVAEPVHLA--YRCRCPQPPGSESVCFEVLGFDIL 338
Qy 268 LVAQLQAWNYK-----ASPEKTYE-----RRELS-----AAY 295
Db 339 LDRKLKPLLEINRAPSGTGDKIDYDKRGVLLNALKLLNIRTSDKRRNLAKQAKAEQR 398
Qy 296 IPFEGGDYYHAATFGGT-----PTHILNLTRECFKGLQDOKKHIDIEAQWHDHSLNKYF 350
Db 399 RLYGONSI---KRLPGSSDWEQHQHLERKEELKERLAQVRKQISREEHENRMHNGY- 454
Qy 351 LFNKPTKILSPE 362
Db 455 -----RRIYPPE 461

RESULT 7
PCT-US02-09944-432
; Sequence 432, Application PC/TUS0209944
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: DAFFO, Abel
; APPLICANT: JONES, Anissa L.
; APPLICANT: TRAN, Alanna-Phung B.
; APPLICANT: DAHL, Christopher R.
; APPLICANT: GIETZEN, Darryl
; APPLICANT: CHINN, Joyce
; APPLICANT: DUFOUR, Gerard E.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: YU, Jimmy Y.
; APPLICANT: TUASON, Olivia
; APPLICANT: YAP, Pierre E.
; APPLICANT: AMSHEY, Stefan R.
; APPLICANT: DAUGHERTY, Sean C.
; APPLICANT: DAM, Tam C.
; APPLICANT: LIU, Tommy F.
; APPLICANT: KLEEFELD, Yael
; APPLICANT: NGUYEN, Duy-Viet An
; APPLICANT: GERSTIN, JR., Edward H.
; APPLICANT: PERALTA, Careyna H.
; APPLICANT: DAVID, Marie H.
; APPLICANT: LEWIS, Samantha A.
; APPLICANT: CHEN, Alice J.
; APPLICANT: PANZER, Scott R.
; APPLICANT: HARRIS, Bernard
; APPLICANT: FLORES, Vincent
; APPLICANT: MARWAHA, Rakesh
; APPLICANT: LO, Audrey
; APPLICANT: LAN, Ruth Y.
; APPLICANT: URASHKA, Michael E.
; TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT
; FILE REFERENCE: PT-1231 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/09944
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: 60/280,067; 60/279,619; 60/280,068; 60/291,280; 60/291,849;
60/291,829; 60/299,428; 60/300,001; 60/299,776
; PRIOR FILING DATE: 2001-03-29; 2001-03-28; 2001-03-29; 2001-05-16; 2001-05-17;
2001-05-17; 2001-06-19; 2001-06-20; 2001-06-20
; NUMBER OF SEQ ID NOS: 792
; SOFTWARE: PERL Program
; SEQ ID NO 432
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: LG:118834.9.orfl:2001MAR30
PCT-US02-09944-432
```

```
Query Match 4.0%; Score 85.5; DB 1; Length 368;
Best Local Similarity 23.6%; Pred. No. 1.8;
Matches 62; Conservative 40; Mismatches 84; Indels 77; Gaps 18;

Qy 152 KLTVCLTVFAVCKYIEHYLEDVLESADMYFVGHVRVI--FYVMIDDTSRMPVHLNPLHSL 210
Db 32 ELWIHLAVVACGNRLLEETLV-MLKSAVLF---SHRKIQHFHFTEDS-----LKEPFDK 80
Qy 211 QV-----FEIRSEKRWODI-----SMRMKKTIGEHLAHIQHEVDVFL 247
Db 81 QLQRPDSDSYTKKFEHRIYPIYTFVSGNPOEKKLKPCAAQRL-----FLPVILKDVDSL 134
Qy 248 FCMDDVQVQDNFGVETLGQLV-----AQLQAWMYKASPEKTYERRELSAAYIPFGEGD 302
Db 135 LYVDTDVLFRL--PVDDIWKLLRLFNSTQLAA---MAPE---HEIPKI-GWYSRFARHP 184
Qy 303 FYHAAIFGGTPTHILNLTTR---ECFKGIL-----OD-----KKIDIEAOWHDESHL 346
Db 185 FYSAGVNSG--VMLMNLTRISTQFKNSMPTGLAWEDMLYPLYQKYKNAITMGDQDLL 242
Qy 347 NKYFLFNKPTKILSPEYC-WDYQ 368
Db 243 NLIYFNPECLYVFP--CQWNYR 263

RESULT 8
US-09-540-209B-8400
; Sequence 8400, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540.209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 8400
; LENGTH: 662
; TYPE: PRT
; ORGANISM: B. fragilis
US-09-540-209B-8400

Query Match 4.0%; Score 85.5; DB 5; Length 662;
Best Local Similarity 20.8%; Pred. No. 4;
Matches 59; Conservative 36; Mismatches 86; Indels 103; Gaps 16;

Qy 164 KYIEHYLEDVLESAD-----MYFMVGHVRVIFYVMIDDTSRMPVHLNPLHSLQVFEIRS 217
Db 54 KEIEKALDFFINQKDSLKIRSIFFLVGN-----MADKYSLTTPANEQDPFHSLNNHIK 107
Qy 218 EKRWDISMRRM-----KT-----IGEHILAHIOHEVDVFL----- 247
Db 108 EKEAWDPCKSRGLGNALDSVYKCTDPPRPKIVRDIEVTGNFLINNVEEAKIHWRTKFF 167
Qy 248 -----FCMDVDVQVQDNFGVETLGQVLAQLQAWMYKASPEKTY-----ERRELSAA 294
Db 168 TECSFDDEFC---EYILPYRIGNESLSA-----WREQACOKESYLLDSDPLDLTKA 216
Qy 295 YIPFGEGDFFYHAAT--FCGTPTHILNLTRECFKGLQDOKKHIDIEAOWHDESHLNKYLFL 352
Db 217 IVQV--SGIYYNAGMSKYPFFPT-----FSEL--DQLH-----WGSCDHLAAAYLTF 258
Qy 353 NKPTKILSPEYCWQYIGLPSDIKSVKVAWQTK-----EYNLVRN 392
Db 259 SLRA-----IGIPSTI-DVVPWANRGGGHAWNVMN 289

RESULT 9
US-10-041-007-39
; Sequence 39, Application US/10041007
; GENERAL INFORMATION:
; APPLICANT: Matsuda, Seichi P.T.
```

APPLICANT: Schepmann, Hala G  
TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase  
FILE REFERENCE: P02081US1  
CURRENT APPLICATION NUMBER: US/10/041,007  
CURRENT FILING DATE: 2002-01-07  
PRIOR APPLICATION NUMBER: US 60/259,881  
PRIOR FILING DATE: 2001-01-05  
NUMBER OF SEQ ID NOS: 41  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 39  
LENGTH: 746  
TYPE: PRT  
ORGANISM: Ginkgo biloba  
US-10-041-007-39

Query Match 3.9%; Score 84; DB 6; Length 746;  
Best Local Similarity 17.6%; Pred. No. 6.5; Mismatches 116; Indels 150; Gaps 17;  
Matches 68; Conservative 53;  
QY 109 PQLWDFNPKNRDPDLTVTPWKAPIVWEGTYDTALLEKYATOKLTVGLTVFAVG----- 163  
DB 35 PQLWDFNPKNRDPDLTVTPWKAPIVWEGTYDTALLEKYATOKLTVGLTVFAVG----- 163  
QY 164 ---KYIEHYLEDLESADMYFMVGHVRIFYVIMDDT-----SR 198  
DB 85 KGVEFVRKHMEEMKDEADNHRPSGFEVFPAMLDKSLGLDLPYHLPFISQIHQKROKK 144  
QY 199 MPVVHLNPLHSLQVFEIRSEK-----RWODISMNR----- 228  
DB 145 LQKIPLNVLNHNQATALLYSLGQDVVDVMDQEIITNLQSRDGSFLSSPASTACVFMHTQNR 204  
QY 229 -----MKTIGEHILAHIQHE-----VDFLCMDVDQVQDNFQVETLGQLV----- 269  
DB 205 CLHFLNVLKSGDGYVPCHPDLDFERLWAVDTVERLGIDRYPKKEI-KESLDVYRYWD 263  
QY 270 AQLOANWYKASP-----EFTYERRELSAAYIPFEGGDF 303  
DB 264 AERGVGWARGNPIPDVDDTAMGLRILRLHGYNVSSDVLENFRDEK-----GDF 311  
QY 304 YYHAAIFGGTPTHTLNLTNR---ECFKGILQDKKHDIKHAOWHDSHLN---KYFLFNKPT 356  
DB 312 FCFAGQTQIGVTNLTNLYRCSQVCFPG-----EKIMEEAKFTTTNHLQNALAKNAFADK-- 365  
QY 357 KILSPCYWDYQIGLPSDIK-SVKVAV 382  
DB 366 -----WAVKDLPGCEVEVAIKYPW 384

RESULT 10  
US-10-041-007-37  
Sequence 37, Application US/10041007  
GENERAL INFORMATION:  
APPLICANT: Matsuda, Seichi P.T.  
APPLICANT: Schepmann, Hala G  
TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase  
FILE REFERENCE: P02081US1  
CURRENT APPLICATION NUMBER: US/10/041,007  
CURRENT FILING DATE: 2002-01-07  
PRIOR APPLICATION NUMBER: US 60/259,881  
PRIOR FILING DATE: 2001-01-05  
NUMBER OF SEQ ID NOS: 41  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 37  
LENGTH: 795  
TYPE: PRT  
ORGANISM: Ginkgo biloba  
US-10-041-007-37

Query Match 3.9%; Score 84; DB 6; Length 795;  
Best Local Similarity 17.6%; Pred. No. 7.1;  
Matches 68; Conservative 53; Mismatches 116; Indels 150; Gaps 17;

QY 109 PQLWDFNPKNRDPDLTVTPWKAPIVWEGTYDTALLEKYATOKLTVGLTVFAVG----- 163  
DB 84 PQLWDFNPKNRDPDLTVTPWKAPIVWEGTYDTALLEKYATOKLTVGLTVFAVG----- 163  
QY 164 ---KYIEHYLEDLESADMYFMVGHVRIFYVIMDDT-----SR 198  
DB 134 KGVEFVRKHMEEMKDEADNHRPSGFEVFPAMLDKSLGLDLPYHLPFISQIHQKROKK 193  
QY 199 MPVVHLNPLHSLQVFEIRSEK-----RWODISMNR----- 228  
DB 194 LQKIPLNVLNHNQATALLYSLGQDVVDVMDQEIITNLQSRDGSFLSSPASTACVFMHTQNR 253  
QY 229 -----MKTIGEHILAHIQHE-----VDFLCMDVDQVQDNFQVETLGQLV----- 269  
DB 254 CLHFLNVLKSGDGYVPCHPDLDFERLWAVDTVERLGIDRYPKKEI-KESLDVYRYWD 312  
QY 270 AQLOANWYKASP-----EFTYERRELSAAYIPFEGGDF 303  
DB 313 AERGVGWARGNPIPDVDDTAMGLRILRLHGYNVSSDVLENFRDEK-----GDF 360  
QY 304 YYHAAIFGGTPTHTLNLTNR---ECFKGILQDKKHDIKHAOWHDSHLN---KYFLFNKPT 356  
DB 361 FCFAGQTQIGVTNLTNLYRCSQVCFPG-----EKIMEEAKFTTTNHLQNALAKNAFADK-- 414  
QY 357 KILSPCYWDYQIGLPSDIK-SVKVAV 382  
DB 415 -----WAVKDLPGCEVEVAIKYPW 433

RESULT 11  
US-10-041-007-35  
Sequence 35, Application US/10041007  
GENERAL INFORMATION:  
APPLICANT: Matsuda, Seichi P.T.  
APPLICANT: Schepmann, Hala G  
TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase  
FILE REFERENCE: P02081US1  
CURRENT APPLICATION NUMBER: US/10/041,007  
CURRENT FILING DATE: 2002-01-07  
PRIOR APPLICATION NUMBER: US 60/259,881  
PRIOR FILING DATE: 2001-01-05  
NUMBER OF SEQ ID NOS: 41  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 35  
LENGTH: 814  
TYPE: PRT  
ORGANISM: Ginkgo biloba  
US-10-041-007-35

Query Match 3.9%; Score 84; DB 6; Length 814;  
Best Local Similarity 17.6%; Pred. No. 7.3;  
Matches 68; Conservative 53; Mismatches 116; Indels 150; Gaps 17;

QY 109 PQLWDFNPKNRDPDLTVTPWKAPIVWEGTYDTALLEKYATOKLTVGLTVFAVG----- 163  
DB 103 PQLWDFNPKNRDPDLTVTPWKAPIVWEGTYDTALLEKYATOKLTVGLTVFAVG----- 163  
QY 164 ---KYIEHYLEDLESADMYFMVGHVRIFYVIMDDT-----SR 198  
DB 153 KGVEFVRKHMEEMKDEADNHRPSGFEVFPAMLDKSLGLDLPYHLPFISQIHQKROKK 212  
QY 199 MPVVHLNPLHSLQVFEIRSEK-----RWODISMNR----- 228  
DB 213 LQKIPLNVLNHNQATALLYSLGQDVVDVMDQEIITNLQSRDGSFLSSPASTACVFMHTQNR 272  
QY 229 -----MKTIGEHILAHIQHE-----VDFLCMDVDQVQDNFQVETLGQLV----- 269  
DB 273 CLHFLNVLKSGDGYVPCHPDLDFERLWAVDTVERLGIDRYPKKEI-KESLDVYRYWD 331  
QY 270 AQLOANWYKASP-----EFTYERRELSAAYIPFEGGDF 303



```
Db 162 PQTLOWILNNQLPD-----GSWGECIEL-AYDRVL-----NTLACLLTLKIWNKGDIQVQ 211
QY 164 ---KYIEHYLEDELESADMYFMVGHRYIFYVMIDDT-----SR 198
Db 212 KGVEFVRKHMEEMKDEADNHRPSCFEVVPAMLDEAKSLGLDLPYHLPFISQIHQKRQKK 271
QY 199 MPVVHLNPLHSLQVFEIRSEK-----RWQDISMR-----228
Db 272 LOKIPLNVLNHNHOTALLYSLGLQDVVDQWQEITNLSRGDSFSLSSPASTACVFMHQNK 331
QY 229 -----MKTIGEHILAHIQHE-----VDFLCMDVDQVFQDNFGVETLGOLV--- 269
Db 332 CLHFLNVLKFGDYVPCHPYPLDLFERLMAVDIVRELGDIDRYFAKEI-KESLOYVYRYWD 390
QY 270 AOLQAWMYKASP-----EKFTYERRELSAAYIPFGEQDF 303
Db 391 AERGVGWARCNPIDVDVDTAMGLRLRLHGVNVSSDVLENFRDEK-----GDF 438
QY 304 YYHAAIFGGTPTTHLNLTR---ECFKILQDKKHIDIAQWHDHSHLN-----KYFLFNKPT 356
Db 439 FCFAGOTQIGVTDNLNLYRCQVCFPG-----EKIMEEAKFTTTHLQNALAKNNAFDK-- 492
QY 357 KILSPEYCDYOIGLPSDIK-SVKVAW 382
Db 493 -----WAVKKDLPGEVEYAIKYPW 511

RESULT 15
US-09-540-209B-9947
; Sequence 9947, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FR
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540, 209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 1044
; SEQ ID NO 9947
; LENGTH: 696
; TYPE: PRT
; ORGANISM: B.fragilis
US-09-540-209B-9947
```

```
Query Match 3.9%; Score 82.5; DB 5; Length 696;
Best Local Similarity 21.5%; Pred. No. 8.2;
Matches 34; Conservative 17; Mismatches 48; Indels 59; Gaps 7;

QY 40 GKVILLMLI---VSTVVVVFWEYVNRIPVEGEN-----69
Db 411 GKQOLIELIRQVHNHPSICFWGLFNLKEVGDNPVEYVKELNLAKQEDPTRPTTSASNQ 470
QY 70 -----RWQK-DWTFPS-----WPKNGTHSYQEDNVGRRKGRNG-----DR 105
Db 471 DGNLNFITENIAMNRYDGYGCTPKTLATFLDRTHKKHPELRIGISEYGAGASYHQQDS 530
QY 106 IEEPOLWDFNPKNRPDLVTPWK----APIVWEGTY 139
Db 531 LKQPSASGWHHPENWQTYYYHWNKIIAERPFVW-GTF 567
```

Search completed: May 10, 2002, 11:09:44  
Job time: 202 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 18:40:43 ; Search time 129.59 Seconds  
(without alignments)  
5418.324 Million cell updates/sec

Title: US-09-863-475A-3  
Perfect score: 1500  
Sequence: 1 CTTCCCTTGTAGACTTTC.....CTTGACACTATTCTTAACCA 1500

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 424068 seqs, 234053524 residues

Total number of hits satisfying chosen parameters: 848136

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_NA\_New.\*

- 1: /cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pna/US06\_NEW\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/pna/US08\_NEW\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/pna/US08\_NEW\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	98.8	6.6	1150	6	US-10-115-123-69
2	85.4	5.7	1193	6	US-10-115-123-119
3	84.2	5.6	305	5	US-09-540-2108-12481
4	37.8	2.5	3186	6	US-10-102-524-1730
5	36.8	2.5	8809	5	US-09-053-375B-710
6	36.2	2.4	790	6	US-10-123-155-204
7	36.2	2.4	837	5	US-09-647-041C-7
8	36.2	2.4	3328	5	US-09-053-375B-494
9	36.2	2.4	1049	6	US-10-123-155-358
10	35.6	2.4	658	6	US-10-102-806-97
11	35.4	2.4	39827	6	US-10-041-018-399
12	35.2	2.3	321	5	US-09-736-968A-80
13	35.2	2.3	547	5	US-09-312-283B-14
14	35.2	2.3	6372	5	US-09-736-968A-1
15	34.6	2.3	605	6	US-10-123-155-160
16	34.2	2.3	933	6	US-10-040-647-29
17	34.2	2.3	1635	1	PCT-US02-09785-73
18	34	2.3	240	5	US-09-975-254-212
19	34	2.3	243	5	US-09-540-2108-6095
20	34	2.3	634	6	US-10-108-580-8
21	33.6	2.2	1726	1	PCT-US02-08992-13
22	33.4	2.2	1184	6	PCT-US02-11682-3
23	33.2	2.2	4500	1	PCT-US02-11682-3
24	33.2	2.2	4500	6	US-10-122-067-3
25	33.2	2.2	6568	1	PCT-US02-11682-1
26	33.2	2.2	6568	6	US-10-122-067-1

27	33.2	2.2	7297	6	US-10-105-299-10267	Sequence 10267, A
28	33.2	2.2	8420	6	US-10-105-299-10266	Sequence 10266, A
29	33	2.2	594	6	US-10-123-155-10	Sequence 10, Appl
30	33	2.2	2288	6	US-10-105-299-2436	Sequence 2436, Ap
31	33	2.2	3406	6	US-10-105-299-13185	Sequence 13185, A
32	33	2.2	3406	6	US-10-105-299-13186	Sequence 13186, A
33	32.4	2.2	897	5	US-09-540-2098-4748	Sequence 4748, Ap
34	32.2	2.1	1659	6	US-10-128-714-2541	Sequence 2541, Ap
35	32.2	2.1	1659	6	US-10-128-714-2541	Sequence 2541, Ap
36	32.2	2.1	2001	6	US-10-128-714-1541	Sequence 7541, Ap
37	32.2	2.1	2001	6	US-10-128-714-1541	Sequence 1541, Ap
38	32.2	2.1	3930	1	PCT-US02-09288-39	Sequence 6541, Ap
39	32.2	2.1	4001	6	US-10-128-714-5541	Sequence 39, Appl
40	32.2	2.1	4001	6	US-10-128-714-5541	Sequence 541, App
41	32	2.1	1580	5	US-09-053-375B-13	Sequence 5541, Ap
42	32	2.1	3939	6	US-10-006-063A-226	Sequence 13, Appl
43	32	2.1	3939	6	US-10-006-117A-226	Sequence 226, App
44	32	2.1	3939	6	US-10-006-130A-226	Sequence 226, App
45	32	2.1	3939	6	US-10-006-172A-226	Sequence 226, App

ALIGNMENTS

RESULT 1

US-10-115-123-69  
; Sequence 69, Application US/10115123  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 94 Human Secreted Proteins  
; FILE REFERENCE: P2029G30APID2  
; CURRENT APPLICATION NUMBER: US/10115,123  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: PCT/US99/13418  
; PRIOR FILING DATE: 1998-06-15  
; PRIOR APPLICATION NUMBER: 60/089,507  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,508  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,509  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,510  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/090,112  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090,113  
; PRIOR FILING DATE: 1998-06-22  
; NUMBER OF SEQ ID NOS: 532  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 69  
; LENGTH: 1150  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-115-123-69

Query Match 6.6%; Score 98.8; DB 6; Length 1150;  
Best Local Similarity 65.5%; Pred No. 4.8e-19;

Matches 203; Conservative 0; Mismatches 57; Indels 50; Gaps 2;

QY	370	CAGGAGAAATAATGAATCTCAAGGAGAAAGTAATCTCTGTGATGATGTTGCTCAACC	429
Db	7	caggagaaaataaggaatgtcaagagaaaagtaattctgtcaatgtgtgtctcaact	66
QY	430	GTGTTCTGCTGTTTGGGAATATGTCAACAG-----	461
Db	67	gtgatcatgtgttttgggaatttatcaacagcacagaaggctcttcttctgtgatatat	126
QY	462	-----AATTCCAGAGTGTGTGAGACAGATGCGAAGGACTGGTGGTTCCTCAAGC	513
Db	127	cactcaaaaaaaccaagaaattgatgcagcagtcgccagaaggctggtggtctctgagc	186
QY	514	TGGTTTAAATGGGACCCCACTTATCAAGAGACAAGCTAGAGGACGGAGAGAAAG	573

```
Db 187 tggtttaacaatggatcccaattatcaacaagggaagacacatagacaaagaaaa 246
QY 574 GGTAGA-----AATGGAGATGCGCATTTGAAGAGCCTCAGCTATGGGACTGGT 619
Db 247 ggaagagagagaccacaaaggaagaaatgacacacagagcttcggtatggagctggt 306
QY 620 TCAATCCAAA 629
Db 307 ttaetccaaa 316

RESULT 2
US-10-115-123-119
; Sequence 119, Application US/10115123
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029G30APID2
; CURRENT APPLICATION NUMBER: US/10/115,123
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: PCT/US99/13418
; PRIOR FILING DATE: 1998-06-15
; PRIOR APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 119
; LENGTH: 1193
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-115-123-119

Query Match 5.7%; Score 85.4; DB 6; Length 1193;
Best Local Similarity 64.1%; Pred. No. 4, 2e-15;
Matches 189; Conservative 0; Mismatches 56; Indels 50; Gaps 2;

QY 385 AATGTCGAAGGAAAAGTAATCCCTGGTGTGATGCTGATTCATCAACCGTGGTTCGCTGTTT 444
Db 68 aatgtcaagaaaagtaattctgtcaatgctgtgtctcaactgtgatcattgtgttt 127
QY 445 TGGGAATATGTCACAG-----AATTCCA 468
Db 128 tgggaattatcacagacacagaaaggtcttcttggatatatacactcaaaaacccca 187
QY 469 GAGGTGGTGAGAACAGATGGCAGAGGACTGGTGGTTCCTCAAGCTGGTTTAAATAATGGG 528
Db 188 gaagttagacagacagctgctcagaagggtggtgttctgagctggtttaacaatggg 247
QY 529 ACCCAGATTTATCAAGAAGACACACGTAGAGACCGAGAGAAAGGCTAG----- 579
Db 248 atccacaattatcaagggggaagacatagacacaaagaaagagagagagacc 307
QY 580 -----AATGGAGATCCATTGAAGACCCCTCAGCTATGGGACTGTTCAATCCAAA 629
Db 308 aaaggaagaaaaatgacacacacagagcttcggtatggagctggtttaatcccaaa 362
```

```
RESULT 3
US-09-540-210B-12481
; Sequence 12481, Application US/09540210B
; GENERAL INFORMATION:
```

```
; APPLICANT: Sellhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE
; FILE REFERENCE: PD-1037 CIP
; CURRENT APPLICATION NUMBER: US/09/540,210B
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 08/972,899
; PRIOR FILING DATE: November 18, 1997
; PRIOR APPLICATION NUMBER: 08/395,244
; PRIOR FILING DATE: February 27, 1995
; PRIOR APPLICATION NUMBER: 08/722,922
; PRIOR FILING DATE: September 27, 1996
; PRIOR APPLICATION NUMBER: 60/005,526
; PRIOR FILING DATE: September 29, 1995
; PRIOR APPLICATION NUMBER: 08/824,029
; PRIOR FILING DATE: March 25, 1997
; PRIOR APPLICATION NUMBER: 60/014,010
; PRIOR FILING DATE: March 25, 1996
; PRIOR APPLICATION NUMBER: 08/826,847
; PRIOR FILING DATE: April 10, 1997
; PRIOR APPLICATION NUMBER: 60/015,533
; PRIOR FILING DATE: April 10, 1996
; PRIOR APPLICATION NUMBER: 08/903,555
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/023,308
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/862,178
; PRIOR FILING DATE: May 22, 1997
; PRIOR APPLICATION NUMBER: 60/018,217
; PRIOR FILING DATE: May 23, 1996
; PRIOR APPLICATION NUMBER: 08/881,589
; PRIOR FILING DATE: June 24, 1997
; PRIOR APPLICATION NUMBER: 60/021,275
; PRIOR FILING DATE: June 25, 1996
; PRIOR APPLICATION NUMBER: 08/903,802
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/023,308
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/905,881
; PRIOR FILING DATE: August 1, 1997
; PRIOR APPLICATION NUMBER: 60/025,204
; PRIOR FILING DATE: August 1, 1996
; PRIOR APPLICATION NUMBER: 08/903,471
; PRIOR FILING DATE: July 30, 1997
; PRIOR APPLICATION NUMBER: 60/025,478
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/903,556
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/025,217
; PRIOR FILING DATE: August 22, 1996
; PRIOR APPLICATION NUMBER: 08/937,142
; PRIOR FILING DATE: September 23, 1997
; PRIOR APPLICATION NUMBER: 60/026,598
; PRIOR FILING DATE: September 24, 1996
; PRIOR APPLICATION NUMBER: 08/960,746
; PRIOR FILING DATE: October 29, 1997
; PRIOR APPLICATION NUMBER: 60/030,144
; PRIOR FILING DATE: October 30, 1996
; PRIOR APPLICATION NUMBER: 08/826,847
; PRIOR FILING DATE: April 10, 1997
; PRIOR APPLICATION NUMBER: 60/015,533
; PRIOR FILING DATE: April 10, 1996
; PRIOR APPLICATION NUMBER: 08/755,524
; PRIOR FILING DATE: November 22, 1996
; PRIOR APPLICATION NUMBER: 60/007,495
; PRIOR FILING DATE: November 22, 1995
; PRIOR APPLICATION NUMBER: 09/021,031
; PRIOR FILING DATE: February 10, 1998
; PRIOR APPLICATION NUMBER: 60/039,325
```

```
; FILE REFERENCE: 210121.572
; CURRENT APPLICATION NUMBER: US/10/102.524
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 1863
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1730
; LENGTH: 3186
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-102-524-1730

Query Match      2.5%; Score 37.8; DB 6; Length 3186;
Best Local Similarity 48.8%; Pred. No. 0.59;
Matches 102; Conservative 0; Mismatches 107; Indels 0; Gaps

Qy   941 AGGATATCAGCATGTCGCCATGAACACATTGGGGAGCACATCTGGCCACATCCAGC 1000
Db   193 aggacatcgcaaggagcaggctggaggagcatcgcggaggagctcgacc 252

Qy   1001 ACAGAGGTGCACCTTCCCTCTTCGTCATCGGAGTGATCAAGTCTTTCAAGACACAACCTTCGGGG 1050
Db   253 gcagcggaggagaacaactacttccagctggagcggagaagatccacaccttcggg 312

Qy   1061 TGGAAACTCTGGGCCAGCTGGTAGACACAGCTCCAGGCTGGTGGTACAAGGCCAGTCCCG 1120
Db   313 agatcacagcaggcagctggaggagaagaagcctgagctgcggaaacaagaccggaga 372

Qy   1121 AGAAGTTCACCTATGAGAGCGGGAAC TG 1149
Db   373 tggagaagccggaggagggcaccagggtg 401

RESULT 5
US-09-053-375B-710/c
; Sequence 710, Application US/09053375B
; GENERAL INFORMATION:
; APPLICANT: Chenchik, Alex
; APPLICANT: Bibilashvili, Robert
; TITLE OF INVENTION: Nucleic Acid Arrays
; FILE REFERENCE: CLON-006
; CURRENT APPLICATION NUMBER: US/09/053,375B
; CURRENT FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 1543
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 710
; LENGTH: 8809
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-053-375B-710

Query Match      2.5%; Score 36.8; DB 5; Length 8809;
Best Local Similarity 80.0%; Pred. No. 1.8;
Matches 56; Conservative 0; Mismatches 12; Indels 2; Gaps

Qy   309 CAAGATCTCATCTCAAGATCCCAAGTCAACAACAAG--TCTTCCATCTCTCAAGATCTGGA 366
Db   8788 CAAGATCTCATACCAGAATCCAGGTCAGAAACTGTATCTCCCAGCCTCCAGATTAGGA 8729

Qy   367 TCACAGGAGA 376
Db   8728 TCACTGGTGA 8719

RESULT 6
US-10-123-155-204/c
; Sequence 204, Application US/10123155
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
```







```

RESULT 15
US-10-123-155-160
: Sequence 160, Application US/10123155
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: DeForge, Laura
: APPLICANT: Desnoyers, Luc
: APPLICANT: Flisaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gottisen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Colin K
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zemin
:

```

Search completed: May 17, 2002, 23:59:16  
Job time: 19113 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2002, 11:07:12 ; Search time 13.25 Seconds  
(without alignments)  
1090.260 Million cell updates/sec

Title: US-09-863-475A-4  
Perfect score: 2136  
Sequence: 1 MITMLQDLHVKNISRSKS.....IKSVKVAQTKYINLRNV 394

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2136	100.0	394	1	GATR_MOUSE
2	1510	70.7	371	1	GATR_PIG
3	1476.5	69.1	368	1	GATR_BOVIN
4	690.5	32.3	354	1	BGAT_HUMAN
5	97	4.5	569	1	GCLL_DROME
6	96.5	4.5	301	1	OMPG_ECOLI
7	95.5	4.5	361	1	Y439_METJA
8	95	4.4	505	1	AGAR_ALTAT
9	91.5	4.3	455	1	SYS_PYRAB
10	91.5	4.3	455	1	SYS_PYRHO
11	91	4.3	323	1	SYK3_PASMU
12	90.5	4.2	496	1	RECO_BACSU
13	90	4.2	1478	1	YAWB_SCHPO
14	88.5	4.1	2896	1	HCYG_OCTDO
15	88	4.1	393	1	KTRL_YEAST
16	88	4.1	534	1	FM02_CAVPO
17	87.5	4.1	1941	1	UBR1_KLOLA
18	87	4.1	865	1	SYV_THEMA
19	87	4.1	1102	1	YM66_YEAST
20	86	4.0	255	1	RL4_PYRAB
21	86	4.0	644	1	YEN9_YEAST
22	86	4.0	1258	1	ATBI_HUMAN
23	86	4.0	1380	1	SP14_YEAST
24	85.5	4.0	887	1	TOP1_BACAN
25	85.5	4.0	1541	1	MRP2_RAT
26	85	4.0	1220	1	ATBI_PIG
27	84.5	4.0	724	1	SECI_YEAST
28	84.5	4.0	730	1	PHLC_PSEAE
29	84.5	4.0	1826	1	SUIS_HUMAN
30	84	3.9	396	1	DCAM_YEAST
31	84	3.9	1258	1	ATBI_RAT
32	83.5	3.9	386	1	CVCB_PEA
33	83.5	3.9	406	1	SBMA_ECOLI

RESULT	ID	GATR_MOUSE	STANDARD	PRT	394 AA
AC	P23336				
DT	01-NOV-1991	(Rel. 20, Created)			
DT	01-NOV-1991	(Rel. 20, Last sequence update)			
DT	30-MAY-2000	(Rel. 39, Last annotation update)			
DE	N-ACETYLGLUCOSAMINIDE ALPHA-1,3-GALACTOSYLTRANSFERASE (EC 2.4.1.151)				
DE	(GALACTOSYLTRANSFERASE) (UDP-GALACTOSE:BETA-D-GALACTOSYL-1,4-N-ACETYL-D-GLUCOSAMINIDE ALPHA-1,3-GALACTOSYLTRANSFERASE)				
GN	GGTA1 OR GGTA-1				
OS	Mus musculus (Mouse)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=90046769; Pubmed=2510162;				
RA	Larsen R.D., Rajan V.P., Ruff M.M., Kukowska-Latallo J.,				
RA	Cummings R.D., Lowe J.B.;				
RT	"Isolation of a cDNA encoding a murine				
RT	UDPgalactose:beta-D-galactosyl-1,4-N-acetyl-D-glucosaminide				
RT	alpha-1,3-galactosyltransferase: expression cloning by gene				
RT	transfer."				
RL	Proc. Natl. Acad. Sci. U.S.A. 86:8227-8231(1989).				
RP	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92184813; Pubmed=1544928;				
RA	Joziasse D.H., Shaper N.L., Kim D., van den Eljnden D.H., Shaper J.H.;				
RT	"Murine alpha 1,3-galactosyltransferase. A single gene locus				
RT	specifies four isoforms of the enzyme by alternative splicing."				
RL	J. Biol. Chem. 267:5534-5541(1992).				
CC	-1- FUNCTION: TRANSFER OF GALACTOSE FROM UDP-GALACTOSE TO AN				
CC	ACCEPTOR MOLECULE (R).				
CC	-1- CATALYTIC ACTIVITY: UDP-GALACTOSE + BETA-D-GALACTOSYL-1,4-				
CC	N-ACETYL-D-GLUCOSAMINYL-R = UDP + ALPHA-D-GALACTOSYL-1,3-				
CC	BETA-D-GALACTOSYL-1,4-N-ACETYL-D-GLUCOSAMINYL-R.				
CC	-1- COFACTOR: MANGANESE.				
CC	-1- PATHWAY: GLYCOSYLATION.				
CC	-1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND				
CC	FORM IN TRANS CISTERNAE OF GOLGI.				
CC	-1- SIMILARITY: STRONG, TO BLOOD GROUP AB TRANSFERASE.				
CC	-1- SIMILARITY: STRUCTURAL SIMILARITY WITH THE OTHER MAMMALIAN				
CC	GLYCOSYLTRANSFERASES.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; M26925; AAA37657.1; -;				
DR	EMBL; M85153; AAA37711.1; -;				
DR	PIR; A34417; A34417.				

DR MGD; MCI:95704; Ggta1.  
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;  
KW Signal-anchor; Golgi stack; Alternative splicing; Manganese.  
FT DOMAIN 1 41  
FT TRANSFERASE 42 60  
FT TRANSFERASE 61 394  
FT TRANSFERASE 83 83  
FT CARBOHYD 319 319  
FT VARSPLIC 1 35  
FT VARSPLIC 62 62  
SQ SEQUENCE 394 AA; 46475 MW; 7766831640D1BBF7 CRC64;

Query Match 100.0%; Score 2136; DB 1; Length 394;  
Best Local Similarity 100.0%; Pred. No. 1.2e-173;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MITMLQDLHVNKISRSKSETSLPSSRSQGEKIMNKGKVLMLLIVSTVVVFWYEV 60  
Db 1 MITMLQDLHVNKISRSKSETSLPSSRSQGEKIMNKGKVLMLLIVSTVVVFWYEV 60  
Qy 61 NRIPEVGNRWQKDWFFSWFKNGTHSQEDNVEGRKGRNGDRIEPQLWDWPNKRN 120  
Db 61 NRIPEVGNRWQKDWFFSWFKNGTHSQEDNVEGRKGRNGDRIEPQLWDWPNKRN 120  
Qy 121 PDVLTVPWKAPIVWEGYDTALLEKYATOKLTGLTVFAVGKYEIHYLEDFLSADMY 180  
Db 121 PDVLTVPWKAPIVWEGYDTALLEKYATOKLTGLTVFAVGKYEIHYLEDFLSADMY 180  
Qy 181 FMVGHVRYFYVIMDDTSRMPVHLNPLSLQVFEIRSEKRWQDISMRMKTIGEHLAHI 240  
Db 181 FMVGHVRYFYVIMDDTSRMPVHLNPLSLQVFEIRSEKRWQDISMRMKTIGEHLAHI 240  
Qy 241 QHEVDFLFCMDVDQVDFQNFVETLGQVLAOLQAWYKASPEKFTYERRELSAAVIFPGE 300  
Db 241 QHEVDFLFCMDVDQVDFQNFVETLGQVLAOLQAWYKASPEKFTYERRELSAAVIFPGE 300  
Qy 301 GDFYHAAIFGTPPHILNLTRFCKGLQDKKHDEAQAQWHDHSLNKFYLFNKPRTKILS 360  
Db 301 GDFYHAAIFGTPPHILNLTRFCKGLQDKKHDEAQAQWHDHSLNKFYLFNKPRTKILS 360  
Qy 361 PEYCDYOIGLPSDIKSVKVAQWTKYENLVNRNV 394  
Db 361 PEYCDYOIGLPSDIKSVKVAQWTKYENLVNRNV 394

RESULT 2  
GATR\_PIG STANDARD; PRT: 371 AA.  
AC P50127;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE N-ACETYLGLYCOSYLTRANSFERASE (UDP-GALACTOSE:BETA-D-GALACTOSYL-1,4-N-ACETYL-  
DE (GALACTOSYLTRANSFERASE) (UDP-GALACTOSE:BETA-D-GALACTOSYL-1,4-N-ACETYL-  
DE D-GLUCOSAMINIDE ALPHA-1,3-GALACTOSYLTRANSFERASE).  
GN GGTAL.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=YORKSHIRE; TISSUE=Spleen;  
RX MEDLINE=95104914; PubMed=7528726;  
RA Strahan K.M., Gu F., Preece A.F., Gustavsson I., Andersson L.,  
RA Gustafsson K.;  
RT "cDNA sequence and chromosome localization of pig alpha 1,3  
RT galactosyltransferase.";  
RL Immunogenetics 41:101-105(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Spleen, and Liver;

RA Sandrin M.S., Dabkowski P.L., Henning M.M., Moutouris E.,  
RA McKenzie I.F.C.;  
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.  
CC FUNCTION: TRANSFER OF GALACTOSE FROM UDP-GALACTOSE TO AN  
CC ACCEPTOR MOLECULE (R).  
CC CATALYTIC ACTIVITY: UDP-GALACTOSE + BETA-D-GALACTOSYL-1,4-  
CC N-ACETYL-D-GLUCOSAMINYL-R = UDP + ALPHA-D-GALACTOSYL-1,3-  
CC BETA-D-GALACTOSYL-1,4-N-ACETYL-D-GLUCOSAMINYL-R.  
CC COFACTOR: MANGANESE.  
CC PATHWAY: GLYCOSYLATION.  
CC SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND  
CC FORM IN TRANS CISTERNAE OF GOLGI.  
CC SIMILARITY: STRONG, TO BLOOD GROUP AB TRANSFERASE.  
CC SIMILARITY: STRUCTURAL SIMILARITY WITH THE OTHER MAMMALIAN  
CC GLYCOSYLTRANSFERASES.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L36152; AAA73558.1; -;  
DR EMBL; L36535; AAA58775.1; -;  
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;  
KW Signal-anchor; Golgi stack; Alternative splicing; Manganese.  
FT DOMAIN 1 6  
FT TRANSFERASE 7 22  
FT TRANSFERASE 23 371  
FT CARBOHYD 59 59  
FT CARBOHYD 296 296  
FT VARSPLIC 27 39  
FT CONFLICT 227 227  
SQ SEQUENCE 371 AA; 43764 MW; CFC715E8D89993D4 CRC64;

Query Match 70.7%; Score 1510; DB 1; Length 371;  
Best Local Similarity 72.7%; Pred. No. 1.2e-120;  
Matches 271; Conservative 48; Mismatches 38; Indels 16; Gaps 4;

Qy 36 MNVKGKVLMLLIVSTVVVFWYEVN-----RIPEVGNRWQKDWFFSWFKN 83  
Db 1 MNVKGKVLMLLIVSTVVVFWYEVNINSPEGLFWYQSKNPEVGSAA-QRGWFFSWFKN 59  
Qy 84 GTHSY--QEDNVEGRKGRNGDRIEPQLWDWPNKRPDVLTVTPWKAPIVWECTYDT 141  
Db 60 GTHSYHEEDAIGNEKEQKEDNRGELP-LVDWPNKRPVTVTPWKAPIVWECTYTNR 118  
Qy 142 ALLEKYATQKLTGLTVFAVGKYEIHYLEDFLSADMYFMVGHVRYFYVIMDDTSRMPV 201  
Db 119 AVLNDVYAKQKLTGLTVFAVGKYEIHYLEDFLSANTYFMVGHVRYFYVIMDDTSRMPV 178  
Qy 202 VHLNPLSLQVFEIRSEKRWQDISMRMKTIGEHLAHIQHEVDFLFCMDVDQVDFQNF 261  
Db 179 IELGLPSFKVFEIRSEKRWQDISMRMKTIGEHLAHIQHEVDFLFCMDVDQVDFQNF 238  
Qy 262 VETLGLVAOLQAWYKASPEKFTYERRELSAAVIFPGEQDFYHAAIFGTPPHILNL 321  
Db 239 VETLGLVAOLQAWYKASPEKFTYERRELSAAVIFPGEQDFYHAAIFGTPPHILNL 298  
Qy 322 RECFCGILQDKKHDEAQAQWHDHSLNKFYLFNKPRTKILSPEYCDYOIGLPSDIKSVKVA 381  
Db 299 QECFCGILQDKKHDEAQAQWHDHSLNKFYLFNKPRTKILSPEYCDYOIGLPSDIKSVKVA 358  
Qy 382 WQTKYENLVNRNV 394  
Db 359 WQTKYENLVNRNV 371

RESULT 3  
GATR\_BOVIN

QY	204	LNP	LSLQVFEIRSEKRWDISMRMKTITGEHILAHQIHEVDFLFCMDVDQVQFQDNFGVE	267
Db	178	LGP	LSRFKFKIKPEKRWQDISMRMKTITGEHIVAHQIHEVDFLFCMDVDQVQFQDKFGVE	237
QY	264	TLG	OLVAQLQAWMYKASPEKFTYERRELSAAYIPGEGDFYHAAIFGCTPTHILNLTRE	323
Db	238	TLG	ESVAQLQAWMYKADPNDFYERRKESAAYIPGEGDFYHAAIFGCTPQVUNITOE	297
QY	324	CFK	GLIQDQKKHDTAEQWHDHSHLNKLYFLFNKPTKLTLSPEYCWQDYQIGLPSDTKSYKVAWQ	383
Db	298	CFK	GILKDKKNDIEAQWHDHSHLNKLYFLNPKTKLTLSPEYCWQDYHIGLPADIKLVKMSQ	357
QY	384	TKE	YNLVRNV 394	
Db	358	TKE	YNVRNV 368	
RESULT 4				
BGAT_HUMAN				
ID	BGAT_HUMAN	STANDARD;	PRT;	354 AA.
AC	PL6442;			
DT	01-AUG-1990 (Rel. 15, Created)			
DT	01-DEC-1992 (Rel. 24, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	HISTO-BLOOD GROUP ABO SYSTEM TRANSFERASE (NAGAT) [INCLUDES:			
DE	GLYCOPROTEIN-FUCOSYL GALACTOSIDE ALPHA-N-			
DE	ACETYLGALACTOSAMINYLTRANSFERASE (EC 2.4.1.40) (FUCOSYLGLYCOPROTEIN			
DE	ALPHA-N-ACETYLGALACTOSAMINYLTRANSFERASE) (HISTO-BLOOD GROUP A			
DE	TRANSFERASE) (A TRANSFERASE); GLYCOPROTEIN-FUCOSYL GALACTOSIDE ALPHA-			
DE	GALACTOSYLTRANSFERASE (EC 2.4.1.37) (FUCOSYLGLYCOPROTEIN 3-ALPHA-			
DE	GALACTOSYLTRANSFERASE) (HISTO-BLOOD GROUP B TRANSFERASE) (B			
DE	TRANSFERASE)]].			
OS	ABO.			
GN	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RP	[1]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RX	MEDLINE=90110098; PubMed=2104828;			
RA	Yamanoto F.-I., Clausen H., White T., Tsuji T., White T., Clausen H.,			
RA	Hakomori S.-I.;			
RT	"Cloning and characterization of DNA complementary to human UDP-			
RT	GalNAc: Fuc alpha 1-->2Gal alpha 1-->3GalNAc transferase (histo-blood			
RT	group A transferase) mRNA.";			
RT	J. Biol. Chem. 265:1146-1151(1990).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90238543; PubMed=2333095;			
RA	Yamanoto F.-I., Clausen H., White T., Marken J., Hakomori S.-I.;			
RT	"Molecular genetic basis of the histo-blood group ABO system.";			
RT	Nature 345:229-233(1990).			
RN	[3]			
RP	CHARACTERIZATION.			
RX	MEDLINE=91035461; PubMed=2121736;			
RA	Yamanoto F.-I., Hakomori S.-I.;			
RT	"Sugar-nucleotide donor specificity of histo-blood group A and B			
RT	transferases is based on amino acid substitutions.";			
RN	J. Biol. Chem. 265:19257-19262(1990).			
CC	-1- FUNCTION: THIS PROTEIN IS THE BASIS OF THE ABO BLOOD GROUP SYSTEM			
CC	THE HISTO-BLOOD GROUP ABO INVOLVES THREE CARBOHYDRATE ANTIGENS: A,			
CC	B, AND H. A, B, AND AB INDIVIDUALS EXPRESS A GLYCOSYLTRANSFERASE			
CC	ACTIVITY THAT CONVERTS THE H ANTIGEN TO THE A ANTIGEN (BY ADDITION			
CC	OF UDP-GALNAc) OR TO THE B ANTIGEN (BY ADDITION OF UDP-GAL),			
CC	WHEREAS O INDIVIDUALS LACK SUCH ACTIVITY.			
CC	-1- CATALYTIC ACTIVITY: UDP-N-ACETYL-D-GALACTOSAMINE + GLYCOPROTEIN			
CC	ALPHA-L-FUCOSYL-(1,2)-D-GALACTOSE = UDP + N-ACETYL-GLYCA-D-			
CC	GALACTOSAMINYL-(1,3)-[ALPHA-L-FUCOSYL-(1,2)]-D-GALACTOSE.			
CC	-1- CATALYTIC ACTIVITY: UDP-GALACTOSE + GLYCOPROTEIN ALPHA-L-FUCOSYL-			
CC	(1,2)-D-GALACTOSE = UDP + GLYCOPROTEIN ALPHA-D-GALACTOSYL-(1,3)-			
CC	[ALPHA-L-FUCOSYL-(1,2)]-D-GALACTOSE.			
CC	-1- PATHWAY: GLYCOSYLATION.			

```
CC FORM IN TRANS CISTERNAE OF GOLGI. SOLUBLE FORM IN BODY FLUIDS.
CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC PROTEOLYTIC PROCESSING.
CC -1- POLYMORPHISM: THE SEQUENCE SHOWN IS THAT OF THE A TRANSFERASE. THE
CC B FORM DIFFERS BY A FEW RESIDUES SUBSTITUTIONS, THE O PHENOTYPE IS
CC RESULT OF A SINGLE BASE FRAMESHIFT DELETION IN THE N-TERMINAL
CC EXTREMITY OF THE GENE.
CC -1- SIMILARITY: STRONG. TO N-ACETYLGLUCOSAMINIDE ALPHA-1,3-
CC GALACTOSYLTRANSFERASE (EC 2.4.1.151).
CC -1- SIMILARITY: STRUCTURAL SIMILARITY WITH THE OTHER MAMMALIAN
CC GLYCOSYLTRANSFERASES.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J05175; AAA36792.1; -.
CC DR PIR; A34933; A34933.
CC DR PIR; S09593; S09593.
CC DR MIM; I10300; -.
CC KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
CC Signal-anchor; Golgi stack; Polymorphism; Blood group antigen.
CC CHAIN 54 354
CC FT FUGOSYLGLYCOPROTEIN ALPHA-N-
CC FT ACETYLGLACTOSAMINYLTRANSFERASE, SOLUBLE
CC FT FORM.
CC FT DOMAIN 1 32
CC FT TRANSHEM 33 53
CC FT CYTOPLASMIC (POTENTIAL).
CC FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
CC FT (POTENTIAL).
CC FT LUMENAL, CATALYTIC (POTENTIAL).
CC FT N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CR -> W.
CC FT /FTID=VAR_003408.
CC FT P -> L (IN ALLELE A2).
CC FT /FTID=VAR_003409.
CC FT R -> G (IN GROUP B TRANSFERASE).
CC FT /FTID=VAR_003410.
CC FT G -> S (IN GROUP B TRANSFERASE).
CC FT /FTID=VAR_003411.
CC FT L -> M (IN GROUP B TRANSFERASE;
CC FT IMPORTANT FOR THE SPECIFICITY).
CC FT /FTID=VAR_003412.
CC FT G -> A (IN GROUP B TRANSFERASE;
CC FT IMPORTANT FOR THE SPECIFICITY).
CC FT /FTID=VAR_003413.
CC FT R -> W (IN ALLELE B3 OF GROUP B
CC FT TRANSFERASE).
CC FT /FTID=VAR_003414.
CC FT SEQUENCE 354 AA; 40934 MW; A03DA16E630C1608 CRC64;
CC -----
Query Match 32.3%; Score 690.5; DB 1; Length 354;
Best Local Similarity 45.8%; Pred. No. 3.3e-51;
Matches 142; Conservative 47; Mismatches 108; Indels 13; Gaps 4;
QY 95 GRREKG-----RNGDTEEPOLWDFNPKN-----RPDLVTVPKAPIVWECTYDTAL 143
DB 45 GSLRGFCMAVRDPDHLQRLVSLPRVVPQKVLTPCKRDVLVTPWLPAPIVWGTFFNIDI 104
QY 144 LEKYATQKLTGLTVFAVGKIEHYLEDLFLESADMYFMVGHRRVIFYVMIDDTSRMPVH 203
DB 105 LNEQFRLONTTIGLTVFAIKKYVA-FLKLFLEAEKHFMVGHRRVHYVFTDPAAPRV 163
QY 204 LNPFLSLQVEIRESEKRWQDSMRMTTIGHILAHQHEVDFLCMDVDVDFQDNFGVE 263
DB 164 LGTGRQLSVLEVRAYKQMDVSMRNMESDFCERRFLSEVDYLVCDVDMEFRDHYGVE 223
QY 264 TLGQLVAQLQAWMYKASPEKFTYERRELSAAYIPFGEGDFYHAAIFGGTPTTHLNLTR 323
DB 224 ILTFLGTLHPGFCSSREAFYERRPQSQAYIPKDEGDFYLLGGFFGGSVQEVQRLTRA 283
```

```
QY 324 CFKGILODKKHDIQAQWHDHSHLNKYLNFKNPKTKILSPYCWQDYQ-IGLPSDIKSVKVAW 382
DB 284 CHQAMVVDQANGIEAVWHDHSHLNKYLNRHKPTKVLSPYELWQDQLGLGMPAVLRKLRTA 343
QY 383 QTKEYNLVRN 392
DB 344 VPKNHQAVRN 353
RESULT 5
GCLL_DROME STANDARD; PRT; 569 AA.
AC Q01820; Q9VAX5;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE GERM CELL-LESS PROTEIN.
GN GCL OR CG8411.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92370679; PubMed=1380406;
RA Jongens T.A., Hay B., Jan L.Y., Jan Y.N.;
RT "The germ cell-less gene product: a posteriorly localized component
RL Cell 70:569-584 (1992).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY;
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananidis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Arif J.F., Aghayani A., An H.-J., Andrews-pfannkuch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclib J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
```

CC -1- FUNCTION: REQUIRED FOR THE SPECIFICATION OF POLE CELLS AND GERM  
CELL FORMATION. MOTHERS WITH REDUCED GLC FUNCTION GIVE RISE TO  
CC STERILE ADULT PROGENY THAT LACK GERM CELLS.  
CC -1- SUBCELLULAR LOCATION: POLE PLASM.  
CC -1- DEVELOPMENTAL STAGE: OOGENESIS AND EMBRYOGENESIS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M97933; AAA28566.1; -.  
DR EMBL: AE003836; AAF59048.1; -.  
DR PIR: A43317; A43317.  
DR FlyBase: FBgn0005695; gcl.  
DR InterPro: IPR000210; BTB\_POZ.  
DR Pfam: PF00651; BTB; 1.  
DR SMART: SM00225; BTB; 1.  
DR PROSITE: PS0097; BTB; 1.  
KW Developmental protein.  
SQ SEQUENCE 569 AA; 65152 MW; A91A670C4B5F83A4 CRC64;  
  
Query Match 4.5%; Score 97; DB 1; Length 569;  
Best Local Similarity 21.3%; Pred. No. 1-2;  
Matches 50; Conservative 42; Mismatches 81; Indels 62; Gaps 10;  
  
QY 96 RREKGRNGDRI---EELQMDWPNKRNPDVLTVP--WKAPIWEGTYDTAL----- 143  
DB 19 RRRKRSTDSGLKDDPAQLDTPPKK-KLLTTTOYIKALFKKEKNSDVAVMALDKVM 77  
QY 144 -LEKYVATQ-----KLTGLTVFAVGKYIEHYLEDLE 175  
DB 78 HLKVLVSQSPYPTWNGTWREAAQNFIOITILDRITVASLDVFGSMYSDEIE--IE 135  
QY 176 SADMYFVAGHRVIFYV--MTDDTSRPVHNLPLHLSLQVPEI-----RSEKRWQD 223  
DB 136 SADVISLATATLFLHLDGIIDKCAEYVNDISPETAQYVEAAQYGVGVKKSTQWFQ 195  
QY 224 ISMRMKTIGEHLAHIQHEVDFLCMDVD-QVFQDNFGVETLGOLVAQQAQWY 277  
DB 196 INLLSIYKQPNLLRHISIELMSALTASPDLYVMQTEFSLYTL-----LRTWTF 244  
  
RESULT 6  
ID OMPG\_ECOLI STANDARD; PRT; 301 AA.  
AC P76045;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE OUTER MEMBRANE PROTEIN G PRECURSOR.  
GN OMPG OR B1319.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 22-35.  
RC STRAIN=K12;  
RX MEDLINE=98389658; PubMed=9721282;  
RA Fajardo D.A., Cheung J., Ito C., Sugawara E., Nikaido H.,  
RA Mlra R.;  
RT "Biochemistry and regulation of a novel Escherichia coli K-12 porin  
protein, OmpG, which produces unusually large channels.";  
RL J. Bacteriol. 180:4452-4459(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12.";  
RL Science 277:1453-1474(1997).  
CC -1- FUNCTION: FORMS CHANNELS FUNCTIONALLY LARGER THAN THOSE OF  
CC CLASSICAL PORINS.  
CC -1- SUBUNIT: MONOMER.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U49400; AAC34720.1; -.  
DR EMBL: AE000230; AAC74401.1; -.  
DR EcoGene: EGI3428; OmpG.  
KW Outer membrane; Transmembrane; Porin; Signal; Complete proteome.  
FT SIGNAL 1 21  
FT CHAIN 22 301 OUTER MEMBRANE PROTEIN G.  
FT TRANSMEM 27 35 POTENTIAL.  
FT TRANSMEM 47 57 POTENTIAL.  
FT TRANSMEM 64 72 POTENTIAL.  
FT TRANSMEM 89 98 POTENTIAL.  
FT TRANSMEM 104 112 POTENTIAL.  
FT TRANSMEM 129 136 POTENTIAL.  
FT TRANSMEM 149 158 POTENTIAL.  
FT TRANSMEM 172 182 POTENTIAL.  
FT TRANSMEM 186 195 POTENTIAL.  
FT TRANSMEM 201 209 POTENTIAL.  
FT TRANSMEM 213 222 POTENTIAL.  
FT TRANSMEM 230 238 POTENTIAL.  
FT TRANSMEM 240 248 POTENTIAL.  
FT TRANSMEM 254 265 POTENTIAL.  
FT TRANSMEM 269 279 POTENTIAL.  
FT TRANSMEM 289 300 POTENTIAL.  
SQ SEQUENCE 301 AA; 34913 MW; 63D1234F3DCCFECF CRC64;  
  
Query Match 4.5%; Score 96.5; DB 1; Length 301;  
Best Local Similarity 22.4%; Pred. No. 0-57;  
Matches 73; Conservative 52; Mismatches 118; Indels 83; Gaps 21;  
  
QY 72 QKDWPFSPFKNGTHSYQEDNVGGRKGRNGDRIEEPQLMDWPNKRNPDVLTVPWKA 131  
DB 24 RNDWHF-----NIGAMYEIENVEG---YGEDMDGLAEPVS--YFNAAN-----GPWRI 66  
QY 132 PIWV--EGTYDTALLEK--YYATOKLTIVGLTVFAVGKYIEHYLEDLEFLESADMYFMV--GH 185  
DB 67 ALAYYQEGPVDYSAGKRGTWFDRELEV-----HY---QPLENDDFSEGLTGGF 112  
QY 186 RVIFYVMID-----DTSRPVHNLPLHLSLQVFEIRSEKRWQDISMMR-----MKTIGEHIL 237  
DB 113 RNYGYHYVDEPGDKTANQWRKLTAPDWDVKLTDDLRENGW--LSMYKFAANDLNTTG-YAD 169  
QY 238 AHIOHEVDFLFCMDVDQVFDNFGVETLGOLVAQQAQWYKASPEKTYERRELSAAYIP 297  
DB 170 TRVETETGLQYTFNETVALRVNYLE-----RGFNMDDSRNNGEFSQEI-RAYLP 219  
QY 298 FGSGDFYHYAAIFGGTPTTHLNLTRCFKGLQDKKHDIAQWIDESHNLKYLFNKPTK 357  
DB 220 LTUGN---HSV---TPYTRIGLDR-----WSNMDWQDDIEREGHD---FNRVGLF----- 260  
QY 358 ILSPEYCWYQIGLPSDYSKVKVAVQ 383  
DB 261 -----GYDFQNGLSVSLE-YAFEWQ 280



CC	EMBL; AJ248287; CAB50241.1; -
DR	HSSP; P34945; 1SER.
DR	InterPro; IPR002106; AA_trna_ligase_II.
DR	InterPro; IPR02314; trna_synt_2b.
DR	InterPro; IPR02317; trna_synt_ser.
DR	InterPro; IPR003364; Seryl_trna_N.
DR	Pfam; PF02403; Seryl_trna_N; 1.
DR	Pfam; PF00587; trna_synt_2b; 1.
DR	PRINTS; PR00981; TRNASYNTHSER.
DR	PROSITE; PS00179; AA_TRNA_LIGASE_II.1; 1.
DR	PROSITE; PS00339; AA_TRNA_LIGASE_II.2; FALSE_NEG.
KW	Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW	complete proteome.
SO	SEQUENCE 455 AA; 53242 MW; DACF960B4B613502 CRC64;

Query Match	4.3%	Score 91.5	DB 1	Length 455
Best Local Similarity	20.1%	Prsd. No. 2.6		
Matches	87	Conservative	62	Mismatches 139; Indels 145; Gaps 25
QY	10	VNKISMSRSKSEVSLPSRSGSQEKINNVGKVTLLMLIVSTVVVVFWVWVNRIP	VEGN 69	
DB	47	INKLRHNRNKIATEI-----GKRKK-----KGPEVEELLAKSKETIVKR	IESLEK--EVEEL 95	
QY	70	RWQKMDFFPSW-FKNGTH-----SYQEDNV-----EGRRK-----GRN	CNR 105	
DB	96	KKKIDYVL--WRLPNITHPSVIGSEEDNVPIRFWCKARWEGHLEFLEOSQ	GK-----149	
QY	106	IEEQLMDWFNPKNRPDLVJVTWPKADIVWEGTYDTALLEK-----YYATQ	KLTV--G 156	
DB	150	-MEYEVLEW-RPKLHVLDLEIL-----GGADFARAARKVSGSRPYLLN	ETVILDLA 198	
QY	157	LTVPFVACKYTE-----HYLEDRLFESADNVPWGHVRVIFYWIDTSM	PVVHLNPL 207	
DB	199	LIRFALDELKKGTPVIPPYVVRFRFVEEGSTTFEDEDVYIKVEGEDLY	LIPFAE--HPL 257	
QY	208	HSLQVFEIRSEKRWQDISMMRMKTIIGHILAHIQHEYDFLFCMDVDQ	VQDFNFGVETLQG 267	
DB	258	AGMHANEILDGK---DLPLLYV-----GISPCKEAGTAGDKTCIERV	HQ 301	
QY	268	LVAOLQAWTYKASPEKFTYERRLSAAYIPFGEGDFYYHAAIFGCT-----	PHIL 318	
DB	302	-----FHKV--EQVYVSRPEES-----WEWHERIIRNABELFOKLE	IPRVV 341	
QY	319	NLTRECF--KGILQDKHDDTEAOWHDESHLNKYLFLFNKPTKILSP	ECWQVQIGLPSDIK 376	
DB	342	NI---CTGDLGYVAAKYIDIAMWPGGGRFR-----EVSASNCTD-----	379	
QY	377	SVKVAWQTKENYL 389		
DB	380	----WQARRLNI 387		
RESULT 10				
SYS_PYRHO	ID	SYS_PYRHO	STANDARD;	PRT; 455 AA.
AC	OS8441;			
DT	15-DEC-1998	(Rel. 37, Created)		
DT	30-MAY-2000	(Rel. 39, Last sequence update)		
DT	20-AUG-2001	(Rel. 40, Last annotation update)		
DE	SERLY-TRNA SYNTHETASE (EC 6.1.1.11) (SERINE--TRNA	LIGASE) (SERRS).		
GN	SRS OR PH0710 OR PHCF014.			
OS	Pyrococcus horikoshii.			
OC	Archaea: Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.			
OX	NCBI_TaxID=53953;			
RP	[1]			
RN	SEQUENCE FROM N.A.			
RC	STRAIN=OT3;			
RX	MSDLINE=98344137; PubMed=9679194;			
RA	Kawarabayasi Y., Sawada M., Horikawa H., Halkawa Y., Hino Y.,			
RA	Yamanoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,			
RA	Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,			
RA	Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,			
RA	Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,			
RA	Masuchi Y., Shizuya H., Kikuchi H.;			
RT	"Complete sequence and gene organization of the genome of a hyper-			
RT	thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";			
RL	DNA Res. 5:55-76(1998).			
CC	-1- CATALYTIC ACTIVITY: ATP + L-SERINE + TRNA(SER) =			
CC	AMP + DIPHOSPHATE + L-SERYL-TRNA(SER).			
CC	-1- SUBCELLULAR LOCATION: CYTOPLASMIC.			
CC	-1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	at the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			



CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; AP000003; BAA29801.1; AUT\_INIT.  
DR InterPro; IPR002106; AA\_trna\_ligase\_II.  
DR InterPro; IPR003364; Seryl\_trna\_N.  
DR InterPro; IPR002314; trna-synt\_2b.  
DR InterPro; IPR002317; trna-synt\_ser.  
DR Pfam; PF02403; Seryl\_trna\_N; 2.  
DR Pfam; PF00587; crna-synt\_2b; 2.  
DR PRINTS; PR00981; TRNASYNTHSER.  
DR PROSITE; PS00179; AA\_TRNA\_LIGASE\_II\_1; 1.  
DR PROSITE; PS00339; AA\_TRNA\_LIGASE\_II\_2; FALSE\_NEG.  
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
KW Complete proteome.  
SQ SEQUENCE 455 AA; 53255 MW; B8D8A418A0651710 CRC64;

Query Match 4.3%; Score 91.5; DB 1; Length 455;  
Best Local Similarity 21.1%; Pred. No. 2.6;  
Matches 91; Conservative 57; Mismatches 141; Indels 143; Gaps 26;

QY 10 VNKISMSKSGTSLPSSRSGSOEKIMNVKGVILLMLIVSTVVVFWFVYVNR-----PE 65  
DB 47 INLRHERNKTAVEI-----GKRRK-----KGEPVDLLAKSR-----EIVKRIGELENE 91  
QY 66 VGENRWQDMFPSP-FKNGTH-----SYQEDNV-----EGRREKGRNGDRIEPPQ-- 110  
DB 92 VEELKKKIDYVL--WRLPNITHPSVPCKDENDNVPFVGKARVKWKGHLERFLEQSOGK 149  
QY 111 -----LWDFNFKNRDVLVTPWKAPIVWEGTYDTALLEK-----YYATOKLTV---GL 157  
DB 150 MEYETLEW-KPKLHVDLEIL-----GGADPARAAKVGSGRFYLLNEIVILDAL 199  
QY 158 TVEAVGKYIE-----HYLEDLESADMYVMVGHVIFVYVIMDDTSRMPVVLHNLPH 208  
DB 200 INFALDRLEIKGFTPIVPPYVVRRECEGTSFEDFVIYKVEDDLYLPTAB-HPLA 258  
QY 209 SLOVFEIRSEKRWQDISMRMKTIGEHILAHQHEVDFLCMDVDQVDFQDFGVTLOQL 268  
DB 259 GWHANEILDGK---DLPLLYV-----GVSPCRKEAGTAGCKTGKIPRVHQ- 301  
QY 269 VAQLQAWYKASPEKFTYERRELSAAYIPFGEGDYFYHAAIFGGT-----PTHILN 319  
DB 302 -----PHKV--EQFVYSRPEES-----WEMHEKIRNAEELFQLEIPYRVVN 342  
QY 320 LTRRCF--KGILQDKKHIDIEAQWHDHSHLNKYFLFNKPTKILSPCYCWDYQIGLPSDIKS 377  
DB 343 I---CTGDLGYVAAKKYDIEA-WMPGQ-----GKFEVVVSANCTD----- 379  
QY 378 VKVAMQTKENYL 389  
DB 380 ----WQARRLNI 387

RESULT 11  
SYK3\_PASMU  
ID SYK3\_PASMU STANDARD; PRT; 323 AA.  
AC P57824;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE PUTATIVE LYSYL-TRNA SYNTHETASE (EC 6.1.1.6) (LYSINE--TRNA LIGASE)  
DE (LYSRS) (GX).  
DE POXA OR PM0202.  
OS Pasteurella multocida.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Pasteurella.  
OX NCBI\_TaxID=747;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PM70;  
RX MEDLINE=21145866; PubMed=11248100;

RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;  
RT \*Complete genomic sequence of Pasteurella multocida pm70.\*;  
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).  
CC -!- FUNCTION: COULD BE A LYSYL-TRNA SYNTHETASE.  
CC -!- CATALYTIC ACTIVITY: ATP + L-LYSINE + TRNA(LYS) = AMP +  
CC PYROPHOSPHATE + L-LYSYL-TRNA(LYS).  
CC -!- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AE006054; AAK02286.1; -  
DR InterPro; IPR002106; AA\_trna\_ligase\_II.  
DR InterPro; IPR002309; trna-synt\_2.  
DR Pfam; PF00152; crna-synt\_2; 1.  
DR PROSITE; PS00179; AA\_TRNA\_LIGASE\_II\_1; 1.  
DR PROSITE; PS00339; AA\_TRNA\_LIGASE\_II\_2; 1.  
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
KW Complete proteome.  
SQ SEQUENCE 323 AA; 36859 MW; 84AFE23F83417888 CRC64;

Query Match 4.3%; Score 91; DB 1; Length 323;  
Best Local Similarity 20.6%; Pred. No. 1.8;  
Matches 42; Conservative 30; Mismatches 80; Indels 52; Gaps 9;

QY 212 VFEIRSEKRWQDISMRMKTIGEHILAHQH-----EVDFLCMDVDQVDFQDNFG 261  
DB 1 MFE---QENWQSPASIEENLLARAKIIAEIRRFETGRLLEVETPVLSEFGVTDVHLSTFN 57  
QY 262 VETLGQLVAQLQAWYKASPEKFTYERRELSAAYIPFGEGDYFYHAAIFGGTPTHILNLT 321  
DB 58 TTFISPTAEKSKALWLSPE---YHMKRLAA-----GSGPIFQLCHVFRNEA----- 104  
QY 322 RECFCGILQDKKHIDIE---AOHWHDHSHLNKYFLFN-----KPTKILSPCYCWD 366  
DB 105 -----GQHNPFTMLEWY-RPHFDMRYRLINEVDLLOQIILQCKPTESLSQFVFPQ 154  
QY 367 VOIGL-PSDIKSVKVAWQTKENYL 389  
DB 155 EVVGLDPLSAEKAELVAKAKQVHL 178

RESULT 12  
RECO\_BACSU  
ID RECO\_BACSU STANDARD; PRT; 496 AA.  
AC P50729;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE ATP-DEPENDENT DNA HELICASE RECO (EC 3.6.1.-) (RECOMBINATION PROTEIN  
DE S).  
GN RECO OR RECS.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168 / MARRBURG;  
RX MEDLINE=96349105; PubMed=8760912;  
RA Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,  
RA Serror P.;  
RT \*Sequence analysis of the Bacillus subtilis chromosome region between  
RT the serA and kds loci cloned in a yeast artificial chromosome.\*;  
RL Microbiology 142:2005-2016(1996).  
RN [2]







KTR1_YEAST	STANDARD;	PRT;	393 AA.
AC	P27810;		
DT	01-AUG-1992 (Rel. 23, Created)		
DT	01-AUG-1992 (Rel. 23, Last sequence update)		
DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	ALPHA-1,2 MANNOSYLTRANSFERASE KTR1 (EC 2.4.1.131).		
GN	KTR1 OR YOR099W OR YOR3189W.		
OS	Saccharomyces cerevisiae (Baker's yeast).		
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.		
OX	NCBI_TaxID=4932;		
ON	{1}		

Db 134 VIPKEHWSFP-EWIDEEKAAQVRKEMGEKRIIYGDSISYRHMCRFESGFFYRHPD---MD 189  
Qy 176 SADMYFMVGHVRV-----IFYVMIDDTSRMPVVLNPLHSLQVFEIRSEKRWQDISM 226  
Db 190 DYDNWVRVPEPIKHLCHCDIDYDFKFMKDNKKYAF-----ISIKYEATPTLWE----- 240  
Qy 227 MRMTTIGEHILAH--IQHEVDFLFCMDVDQ-----VFQDNFGVETLGQLVAQLQAWMY 277  
Db 241 -----TTRKMEAHPELIHENNMNLDVSDDOGLSYNLCHFWSNFE-----IAALDLWRS 289  
Qy 278 KASPEKFTYERRELSAAYIPGE-----GDF-YYHAAIFGGTPTH 316  
Db 290 PAYSAYFDYLDREGGFFYERWGDAPVHSIGAAFLDRSEIHHFGDIGYHYVPHSHCPIDT 349  
Qy 317 ILNLTRECFCGILQDKKHHDIEAQWHDSEHLNKYFLFNKPTKILSPEYCWQYQIG 370  
Db 350 SIRLANKC-----DCDPSKDFTWHSYSCTTKFYNNKLPKPAG-----WQNHIG 393

Search completed: May 10, 2002, 11:10:36  
Job time: 204 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 16:06:15 ; Search time 2776.7 Seconds  
(without alignments)  
11304.716 Million cell updates/sec

Title: US-09-863-475A-3  
Perfect score: 1500  
Sequence: 1 CCTCCCTTGACTCTTC.....CTTGACACTATTCTTAACCA 1500

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*

1: gb\_ba:\*

2: gb\_htg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vl:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pat:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_sts:\*

28: em\_un:\*

29: em\_vl:\*

30: em\_htg\_hum:\*

31: em\_htg\_inv:\*

32: em\_htg\_other:\*

33: em\_htg\_inv:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	1500	100.0	1500	6	AR075307	AR075307 Sequence
2	1500	100.0	1500	6	AX306447	AX306447 Sequence
3	1500	100.0	1500	6	I34188	I34188 Sequence 3
4	1500	100.0	1500	10	MUSGALA	M26925 Mouse galac
5	1434.8	95.7	3537	6	AX128259	AX128259 Sequence
6	1434.8	95.7	3537	10	AF297615	AF297615 Mus muscu
7	1419.6	94.6	3450	10	MUSGLYTN	M85153 M.musculus
8	1408.4	93.9	3450	6	AR066335	AR066335 Sequence
9	1326.4	88.4	3315	10	BC006810	BC006810 Mus muscu
10	726.6	48.4	1343	4	PIC13G	L36535 Sus scrofa
11	726.6	48.4	1423	6	AR048516	AR048516 Sequence
12	725.2	48.3	3200	10	AF297606S9	AF297614 Mus muscu
13	725.2	48.3	3240	6	AX128258	AX128258 Sequence
14	706	47.1	1092	6	A69346	A69346 Sequence 3
15	706	47.1	1092	6	A72023	A72023 Sequence 3
16	693.8	46.3	1131	6	AX306443	AX306443 Sequence
17	693.8	46.3	1131	9	S71333	S71333 alpha 1,3 g
18	684.2	45.6	1269	4	P1GGGTA	L36152 Sus scrofa
19	684.2	45.6	1269	6	AR119910	AR119910 Sequence
20	684.2	45.6	1269	6	AX306445	AX306445 Sequence
21	684.2	45.6	1412	6	AR066333	AR066333 Sequence
22	684.2	45.6	3640	4	AF221509	AF221509 Sus scrof
23	684.2	45.6	3684	4	AF221508	AF221508 Sus scrof
24	664.6	44.3	1128	6	A69344	A69344 Sequence 1
25	664.6	44.3	1128	6	A72021	A72021 Sequence 1
26	645.6	43.0	1303	6	AX306439	AX306439 Sequence
27	642.4	42.8	1617	6	AX306441	AX306441 Sequence
28	642.4	42.8	1828	4	BOVGSTA	J04989 Bovine alph
29	642.4	42.8	1828	6	AR066334	AR066334 Sequence
30	632.8	42.2	1065	6	A69348	A69348 Sequence 5
31	632.8	42.2	1065	6	A72025	A72025 Sequence 5
32	617.2	41.1	1029	6	A69350	A69350 Sequence 7
33	617.2	41.1	1029	6	A72027	A72027 Sequence 7
34	583.2	38.9	1353	6	AR048515	AR048515 Sequence
35	563.4	37.6	106400	9	AC012518	AC012518 Homo sapi
36	561.8	37.5	1885	6	AX306449	AX306449 Sequence
37	561.8	37.5	1885	9	HUMHGT2	M60263 Human alpha
38	557.6	37.2	1053	9	AF197007	AF197007 Pan trogl
39	529.2	35.3	1048	9	AF196986	AF196986 Homo sapi
40	518.2	34.5	2964	6	AX128247	AX128247 Sequence
41	518.2	34.5	3900	4	AF221510S8	AF221517 Sus scrof
42	516.2	34.4	3322	6	AX128266	AX128266 Sequence
43	511.2	34.1	3745	6	AX128267	AX128267 Sequence
44	488.8	32.6	150402	2	AC078879	AC078879 Homo sapi
45	479	31.9	3135	6	AX128260	AX128260 Sequence

ALIGNMENTS

RESULT	1	AR075307	1500 bp	DNA	linear	PAT 28-AUG-2000
LOCUS	AR075307	Sequence 3 from patent US 5955347.				
DEFINITION	AR075307	Sequence 3 from patent US 5955347.				
ACCESSION	AR075307	Sequence 3 from patent US 5955347.				
VERSION	AR075307.1	GI:10002057				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 1500)					
AUTHORS	Lowe,J.B.					
TITLE	Methods and products for the synthesis of oligosaccharide structures on glycoproteins, glycolipids, or as free molecules, and for the isolation of cloned genetic sequences that determine these structures					
JOURNAL	Patent: US 5955347-A 3 21-SEP-1999;					
FEATURES	Location/Qualifiers					
source	1..1500					
BASE COUNT	406 a 349 c 374 g 371 t					
ORIGIN	/organism="unknown"					

Query Match		100.0%;	Score 1500;	DB 6;	Length 1500;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 1500;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CCTTCCCTTGTAGACTCTTCTTGAATGAGAGTACCGATTCTGCTGAAGACCTCGGGCT	60		
Db	1	CCTTCCCTTGTAGACTCTTCTTGAATGAGAGTACCGATTCTGCTGAAGACCTCGGGCT	60		
Qy	61	CTCAGGCTCTGGGAGTTGGAAACCTGTACCTTCTTCCCTGCTGAGCCCTGCTCCTT	120		
Db	61	CTCAGGCTCTGGGAGTTGGAAACCTGTACCTTCTTCCCTGCTGAGCCCTGCTCCTT	120		
Qy	121	AGGCAGCCAGAGCTGCACAGAACTCGGTTGCTTGTGTTGCTTGGAGGGAACACAG	180		
Db	121	AGGCAGCCAGAGCTGCACAGAACTCGGTTGCTTGTGTTGCTTGGAGGGAACACAG	180		
Qy	181	CTGACGATGAGGCTGACTTTGAACCTCAAGAGATCTGTTACCCAGTCTCTGCTGAATTA	240		
Db	181	CTGACGATGAGGCTGACTTTGAACCTCAAGAGATCTGTTACCCAGTCTCTGCTGAATTA	240		
Qy	241	AGGCCTGTACTACATTTGCTGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCTC	300		
Db	241	AGGCCTGTACTACATTTGCTGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCTC	300		
Qy	301	CATCTCAACAAGATCTCCATGTCAAGATCCAAAGTCAGAAACAACTCTTCCATCTCAAGA	360		
Db	301	CATCTCAACAAGATCTCCATGTCAAGATCCAAAGTCAGAAACAACTCTTCCATCTCAAGA	360		
Qy	361	TCTGGATCACAGGAGAAATAAATGAATGTCNAGGAAAGTAATCCTGTGATGCTGATT	420		
Db	361	TCTGGATCACAGGAGAAATAAATGAATGTCNAGGAAAGTAATCCTGTGATGCTGATT	420		
Qy	421	GTCTCAACCGTGTGCTGCTTTTGGGAATATCTCAACAGAATTCAGAGGTTGGTGAG	480		
Db	421	GTCTCAACCGTGTGCTGCTTTTGGGAATATCTCAACAGAATTCAGAGGTTGGTGAG	480		
Qy	481	AACAGATGGCAGAAAGGACTGCTGCTTCCCAAGCTGGTTTAAAAATGGGACCCACAGTAT	540		
Db	481	AACAGATGGCAGAAAGGACTGCTGCTTCCCAAGCTGGTTTAAAAATGGGACCCACAGTAT	540		
Qy	541	CAAGAACAACACGTAGAAAGCAGAGAGAGAAAGGTAGAAATGAGATCGCATTTGAAGAG	600		
Db	541	CAAGAACAACACGTAGAAAGCAGAGAGAGAAAGGTAGAAATGAGATCGCATTTGAAGAG	600		
Qy	601	CCTCAGCTATGGGACTGCTTCAATCCAAAGAACCCCGCGGATGTTTACAGTGCACCCCG	660		
Db	601	CCTCAGCTATGGGACTGCTTCAATCCAAAGAACCCCGCGGATGTTTACAGTGCACCCCG	660		
Qy	661	TGGAAGCGCGGATGCTGCTGGGAAGGCACTTATGACACAGCTCTGCTGGAAAAGTACTAC	720		
Db	661	TGGAAGCGCGGATGCTGCTGGGAAGGCACTTATGACACAGCTCTGCTGGAAAAGTACTAC	720		
Qy	721	GCCACACAGAAACTCACTGTGGGCTGACAGTGTGCTGTGGCAAAAGTACATTTGAGCAT	780		
Db	721	GCCACACAGAAACTCACTGTGGGCTGACAGTGTGCTGTGGCAAAAGTACATTTGAGCAT	780		
Qy	781	TACTTAGAAGACTTCTCGAGTCTGCTGACATGTACTTTCATGTTGGCCATCGGGTCATA	840		
Db	781	TACTTAGAAGACTTCTCGAGTCTGCTGACATGTACTTTCATGTTGGCCATCGGGTCATA	840		
Qy	841	TTTTACGTATGATAGACACACCTCCCGGATGCTGTGCTGACCTCTGAAACCTCTACAT	900		
Db	841	TTTTACGTATGATAGACACACCTCCCGGATGCTGTGCTGACCTCTGAAACCTCTACAT	900		
Qy	901	TCCTTACAAGTCTTTGATATCAGTCTGTAGAGAGGTGGCAGGATATCAGCATATGCGC	960		
Db	901	TCCTTACAAGTCTTTGATATCAGTCTGTAGAGAGGTGGCAGGATATCAGCATATGCGC	960		
Qy	961	ATGAAGACCATTTGGGAGCAGCATCTCTGGCCACATCCAGCAGAGGTCGACTTCTCTTC	1020		
Db	961	ATGAAGACCATTTGGGAGCAGCATCTCTGGCCACATCCAGCAGAGGTCGACTTCTCTTC	1020		

Qy	1021	TGCATGACGCTGGATCAAGTCTTTCAAGACAACCTTCGGGGTGAACACTCTCGGCCAGCTG	1080		
Db	1021	TGCATGACGCTGGATCAAGTCTTTCAAGACAACCTTCGGGGTGAACACTCTCGGCCAGCTG	1080		
Qy	1081	GTACACAGCTCCAGGCTGCTGTACAGGCCAGTCCCGAGAAGTTCCACCTATGAGAG	1140		
Db	1081	GTACACAGCTCCAGGCTGCTGTGTACAGGCCAGTCCCGAGAAGTTCCACCTATGAGAG	1140		
Qy	1141	CGGAACCTGTCGGCGCGGTACATTTCCATTTCGAGAGGGGATTTTACTACACCGCGCC	1200		
Db	1141	CGGAACCTGTCGGCGCGGTACATTTCCATTTCGAGAGGGGATTTTACTACACCGCGCC	1200		
Qy	1201	ATTTTGGAGAACCCCTACTCACATTTCTCAACCTCACAGGAGTGTCTTTAAGGGATC	1260		
Db	1201	ATTTTGGAGAACCCCTACTCACATTTCTCAACCTCACAGGAGTGTCTTTAAGGGATC	1260		
Qy	1261	CTCAGGACAAGAACATGACATAGAACCCAGTGGCATGATGAGAGCCACTCAACAA	1320		
Db	1261	CTCAGGACAAGAACATGACATAGAACCCAGTGGCATGATGAGAGCCACTCAACAA	1320		
Qy	1321	TACTTCTCTTTCAACAACCCACTAAATCTCTATCTCCAGAGTATTGCTGGGACTATCAG	1380		
Db	1321	TACTTCTCTTTCAACAACCCACTAAATCTCTATCTCCAGAGTATTGCTGGGACTATCAG	1380		
Qy	1381	ATAGCCTGCTCCATGATATTTAAAGTGTCAAGGTAGCTTGGCAGACAAAAGATTAAT	1440		
Db	1381	ATAGCCTGCTCCATGATATTTAAAGTGTCAAGGTAGCTTGGCAGACAAAAGATTAAT	1440		
Qy	1441	TTGCTTAGAATAATGCTGACTTCAAAATTTGATGGAACTTGACACTATTCTAAACA	1500		
Db	1441	TTGCTTAGAATAATGCTGACTTCAAAATTTGATGGAACTTGACACTATTCTAAACA	1500		
RESULT 2					
AX306447	AX306447	Sequence 9 from Patent WO0188096.	1500 bp	DNA	l linear PAT 11-DEC-2001
LOCUS	AX306447				
DEFINITION	AX306447				
ACCESSION	AX306447				
VERSION	AX306447.1	GI:17645668			
KEYWORDS					
SOURCE	Musca sp.				
ORGANISM	Musca sp.				
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;				
	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;				
	Muscomorpha; Muscoidea; Muscidae; Musca.				
REFERENCE	1 (sites)				
AUTHORS	Denning, C. and Clark, J.				
TITLE	Animal tissue for xenotransplantation				
JOURNAL	Patent: WO 0188096-A 9 22-NOV-2001;				
FEATURES	Geron Corporation (US)				
source	Location/Qualifiers				
	1..1500				
	/organism="Musca sp."				
	/db_xref="taxon:30058"				
	277..1461				
	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CA019042.1"				
	/db_xref="GI:17645668"				
	/translation="MTHMLQDLHVNKLSMSRSKSETSLPSSRSQSQEKIMNVKCKVIL				
	LMILIVSTVVVFWVEYVNRIPVEGVENRQKQWPFPSFKNGTSTQEDNVGRRKGRN				
	GDRIEPLQDWDFNPKRPDLVTPNKPAPIVWEGTYDTALLESKYATQKLTGLTVF				
	AVKYIEHYLEDFLESADMYFMVGHVIFVIMDDTSRMPVHLNPLHLSLQVFIRES				
	KRWQDSMMRMKTIQHEVLDVFCMDVDFQDFQDFQDFQDFQDFQDFQDFQDFQDFQ				
	YKASPEKFTYVERRELSAAYTPFGGDFYHAAIFGGTPTHLINLTRECFKGLQDKKH				
	DICAQHDESHLNKYFLFNKPTKILSPYCDYQIGLPSDIKSKYKVAWQTEYNLVRN				
	NV"				
BASE COUNT	406 a	349 c	374 g	371 t	
ORIGIN					
Query Match	100.0%;	Score 1500;	DB 6;	Length 1500;	
Best Local Similarity	100.0%;	Pred. No. 0;			

Matches 1500: Conservative 0: Mismatches 0: Indels 0: Gaps 0:			
QY	1	CCTTCCCTTGTAGACTCTTCTTGGAAATGAGAACTACCGATTCTGCTGAAGACCTCCGGCT	60
Db	1	CCTTCCCTTGTAGACTCTTCTTGGAAATGAGAACTACCGATTCTGCTGAAGACCTCCGGCT	60
QY	61	CTCAGGCTCTGGAGTGGAAACCTGTACCTTCCCTTTCCTCTGCTGAGCCCTGCCCTT	120
Db	61	CTCAGGCTCTGGAGTGGAAACCTGTACCTTCCCTTTCCTCTGCTGAGCCCTGCCCTT	120
QY	121	AGGCAGCCAGAGCTCGACAGAACTCGGTTGCTTTCCTCTTTCCTTGGAGGGAACACAG	180
Db	121	AGGCAGCCAGAGCTCGACAGAACTCGGTTGCTTTCCTCTTTCCTTGGAGGGAACACAG	180
QY	181	CTGACGATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCACGCTCCTTGAATTA	240
Db	181	CTGACGATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCACGCTCCTTGAATTA	240
QY	241	AGGCCTGTACTACATTTGCTTGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCTC	300
Db	241	AGGCCTGTACTACATTTGCTTGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCTC	300
QY	301	CATGTCAACAAGATCTCCATGTCAAGATCCAAGTCAAGAACAAAGTCTTCCATCTCAAGA	360
Db	301	CATGTCAACAAGATCTCCATGTCAAGATCCAAGTCAAGAACAAAGTCTTCCATCTCAAGA	360
QY	361	TCGTGATCAGAGAGAAATTAATGAATGTCAAGGGAAGTAACTCTGTTGATGCTGATT	420
Db	361	TCGTGATCAGAGAGAAATTAATGAATGTCAAGGGAAGTAACTCTGTTGATGCTGATT	420
QY	421	GTCTCAACCGTGTGCTGCTGTTTGGGAATATGTCAACAGAAATTCAGAGGTTGGTGAG	480
Db	421	GTCTCAACCGTGTGCTGCTGTTTGGGAATATGTCAACAGAAATTCAGAGGTTGGTGAG	480
QY	481	AACAGATGCGAAGGACTGGTGGTCCCAAGCTGTTTAAATAATGGACCCACACTAT	540
Db	481	AACAGATGCGAAGGACTGGTGGTCCCAAGCTGTTTAAATAATGGACCCACACTAT	540
QY	541	CAAGAAGACAACTAGAGGACGAGAGAAAGGTTAGAAATGGAGATCGCATTTGAAGAG	600
Db	541	CAAGAAGACAACTAGAGGACGAGAGAAAGGTTAGAAATGGAGATCGCATTTGAAGAG	600
QY	601	CCTCAGCTATGGACTGGTTCAATCAAGAAGCCGCCGGATGTTTGACAGTGACCCCG	660
Db	601	CCTCAGCTATGGACTGGTTCAATCAAGAAGCCGCCGGATGTTTGACAGTGACCCCG	660
QY	661	TGGAAGCGCGGATGTGTGGGAAGGCACTTATGACACAGCTCTGCTGGAAAAAGTACTAC	720
Db	661	TGGAAGCGCGGATGTGTGGGAAGGCACTTATGACACAGCTCTGCTGGAAAAAGTACTAC	720
QY	721	GCCACAGAAACTACTGTGGGCTGACAGTGTGCTGTGGGAAAGTACATTTGAGCAT	780
Db	721	GCCACAGAAACTACTGTGGGCTGACAGTGTGCTGTGGGAAAGTACATTTGAGCAT	780
QY	781	TACTTTAGAGACTTTCTGGAGTCTGCTGACATGACTTATGTTGGCCATCCGGTGCATA	840
Db	781	TACTTTAGAGACTTTCTGGAGTCTGCTGACATGACTTATGTTGGCCATCCGGTGCATA	840
QY	841	TTTTTACGTCATGATAGACACACTCCCGATGCTGCTGACCTGAAACCTCTTACAT	900
Db	841	TTTTTACGTCATGATAGACACACTCCCGATGCTGCTGACCTGAAACCTCTTACAT	900
QY	901	TCCTTACAAGTCTTTGAGATCAGGCTTGAGAGAGGTGGCAGGATATCAGCATGATGCCG	960
Db	901	TCCTTACAAGTCTTTGAGATCAGGCTTGAGAGAGGTGGCAGGATATCAGCATGATGCCG	960
QY	961	ATGAAGACCAATTTGGGAGCACATCTTGGGCCACATCCAGACAGAGTTCGACTTCTTTC	1020
Db	961	ATGAAGACCAATTTGGGAGCACATCTTGGGCCACATCCAGACAGAGTTCGACTTCTTTC	1020
QY	1021	TGCATGGAGCTGATCAAGTCTTTCAAGACAACTTGGGGTGGAAACTCTGGGCCAGCTG	1080
Db	1021	TGCATGGAGCTGATCAAGTCTTTCAAGACAACTTGGGGTGGAAACTCTGGGCCAGCTG	1080

QY	1081	GTAGCAGCTCCAGGCCCTGGTGTACAAAGGCCAGTCCCGAGAAGTTCACTATGAGAG	1140
Db	1081	GTAGCAGCTCCAGGCCCTGGTGTACAAAGGCCAGTCCCGAGAAGTTCACTATGAGAG	1140
QY	1141	CGGGAAGTCTGCGCCCGGTACATTTCCATTTCGGAGAGGGGATTTTACTACACGGGCC	1200
Db	1141	CGGGAAGTCTGCGCCCGGTACATTTCCATTTCGGAGAGGGGATTTTACTACACGGGCC	1200
QY	1201	ATTTTGGAGGAACGCTACTACATTTCAACCTTCACACGGGAGTGTCTTAAGGGGATC	1260
Db	1201	ATTTTGGAGGAACGCTACTACATTTCAACCTTCACACGGGAGTGTCTTAAGGGGATC	1260
QY	1261	CTCCAGGACAAGAACATGACATAGAACCCAGTGGCATGATGAGAGCCACCTCAACAAA	1320
Db	1261	CTCCAGGACAAGAACATGACATAGAACCCAGTGGCATGATGAGAGCCACCTCAACAAA	1320
QY	1321	TACTTCTCTTTCAACAAACCCACTAAATTCCTATCTCCAGAGTATTTGCTGGGACTATCAG	1380
Db	1321	TACTTCTCTTTCAACAAACCCACTAAATTCCTATCTCCAGAGTATTTGCTGGGACTATCAG	1380
QY	1381	ATAGGCTGCTCCTTCAGATATTAAGTGTCAAGGTAGCTTGGCAGACAAAAGAGTATAAT	1440
Db	1381	ATAGGCTGCTCCTTCAGATATTAAGTGTCAAGGTAGCTTGGCAGACAAAAGAGTATAAT	1440
QY	1441	TTGGTTAGAAATAATGCTGACTTCAAAATGTGATGGAACCTTGACACTATTTCTAACCA	1500
Db	1441	TTGGTTAGAAATAATGCTGACTTCAAAATGTGATGGAACCTTGACACTATTTCTAACCA	1500
RESULT 3			
134188			
LOCUS			
DEFINITION			
Sequence 3 from patent US 5595900.			
ACCESSION			
134188			
VERSION			
134188.1 GI:1824979			
KEYWORDS			
Unknown.			
SOURCE			
ORGANISM			
REFERENCE			
1 (bases 1 to 1500)			
AUTHORS			
Lowe,J.B.			
TITLE			
Methods and products for the synthesis of oligosaccharide structures on glycoproteins, glycolipids, or as free molecules, and for the isolation of cloned genetic sequences that determine these structures			
JOURNAL			
Patent: US 5595900-A 3 21-JAN-1997;			
FEATURES			
source			
Location/Qualifiers			
1..1500			
/organism="unknown"			
BASE COUNT			
406 a 349 c 374 g 371 t			
ORIGIN			
Query Match			
100.0%; Score 1500; DB 6; Length 1500;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	CCTTCCCTTGTAGACTCTTCTTGGAAATGAGAACTACCGATTCTGCTCAACAGCCCTCCGGCT	60
Db	1	CCTTCCCTTGTAGACTCTTCTTGGAAATGAGAACTACCGATTCTGCTCAACAGCCCTCCGGCT	60
QY	61	CTCAGGCTCTGGAGTGGAAACCTGTACCTTCCCTTTCCTCTGCTGAGCCCTGCCCTT	120
Db	61	CTCAGGCTCTGGAGTGGAAACCTGTACCTTCCCTTTCCTCTGCTGAGCCCTGCCCTT	120
QY	61	CTCAGGCTCTGGAGTGGAAACCTGTACCTTCCCTTTCCTCTGCTGAGCCCTGCCCTT	120
Db	61	CTCAGGCTCTGGAGTGGAAACCTGTACCTTCCCTTTCCTCTGCTGAGCCCTGCCCTT	120
QY	121	AGGCAGCCAGAGCTCGACAGAACTCGGTTGCTTTCCTCTTTCCTTGGAGGGAACACAG	180
Db	121	AGGCAGCCAGAGCTCGACAGAACTCGGTTGCTTTCCTCTTTCCTTGGAGGGAACACAG	180
QY	181	CTGACCATGAGGCTGACTTTGAACATCAAGAGATCTGCTTACCCACGCTCCTTGAATTA	240
Db	181	CTGACCATGAGGCTGACTTTGAACATCAAGAGATCTGCTTACCCACGCTCCTTGAATTA	240

Qy 241 AGGCGTGACTACATTTGCCCTGGACCTAAGATTTTTCATGATCACTATGCTTTCAGATCTC 300  
Db 241 AGGCGTGACTACATTTGCCCTGGACCTAAGATTTTTCATGATCACTATGCTTTCAGATCTC 300  
Qy 301 CATGTCACACAGATCTCCATGTCAGATCCAAAGTCAGAAACAGTCTTCCATCTCAAGA 360  
Db 301 CATGTCACACAGATCTCCATGTCAGATCCAAAGTCAGAAACAGTCTTCCATCTCAAGA 360  
Qy 361 TCTGGATCACAGGAGAAAATTAATGATGTCAGGGAAGAAATTAATCCTGTTGATGCTGATT 420  
Db 361 TCTGGATCACAGGAGAAAATTAATGATGTCAGGGAAGAAATTAATCCTGTTGATGCTGATT 420  
Qy 421 GTCTCAACCGTGTGCTGTTTGGGAATATGTCACAGAAATTCACAGAGTTGGTGAG 480  
Db 421 GTCTCAACCGTGTGCTGTTTGGGAATATGTCACAGAAATTCACAGAGTTGGTGAG 480  
Qy 481 AACAGATGGCAGAGGACTGTGCTTCCCAAGCTGGTTTAAANAATGGACCCACAGTTAT 540  
Db 481 AACAGATGGCAGAGGACTGTGCTTCCCAAGCTGGTTTAAANAATGGACCCACAGTTAT 540  
Qy 541 CAAGAACAACAGTGAAGAGCAGAGAGAAAGGTTAGAAATGGAGATCGCATTTGAAGAG 600  
Db 541 CAAGAACAACAGTGAAGAGCAGAGAGAAAGGTTAGAAATGGAGATCGCATTTGAAGAG 600  
Qy 601 CCTCAGCTATGGGACTGGTTCAATCCAAAGAACCCCGCGGATGTTTTCAGAGTCAACCCG 660  
Db 601 CCTCAGCTATGGGACTGGTTCAATCCAAAGAACCCCGCGGATGTTTTCAGAGTCAACCCG 660  
Qy 661 TGGAGGCGCGGATGTTGGGAGGACCTTATGACACAGCTCTGCTGGAAAAGTACTAC 720  
Db 661 TGGAGGCGCGGATGTTGGGAGGACCTTATGACACAGCTCTGCTGGAAAAGTACTAC 720  
Qy 721 GCCACACAGAAATCACTGTGGGCTGACAGTGTGCTGTGGAAAAGTACATTTGAGCAT 780  
Db 721 GCCACACAGAAATCACTGTGGGCTGACAGTGTGCTGTGGAAAAGTACATTTGAGCAT 780  
Qy 781 TACTTAGAAGACTTTCGGAGTCTGCTGACATGATCTTTCATGTTGGCCATCGGGTCATA 840  
Db 781 TACTTAGAAGACTTTCGGAGTCTGCTGACATGATCTTTCATGTTGGCCATCGGGTCATA 840  
Qy 841 TTTTACGTCATGATAGACGACACCTCCCGGATGCTGTGCTGCACTTGAACCCCTCTACAT 900  
Db 841 TTTTACGTCATGATAGACGACACCTCCCGGATGCTGTGCTGCACTTGAACCCCTCTACAT 900  
Qy 901 TCCTTACAAGTCTTTGAGATCAGTCTGAGAAAGAGTGGCAGGATATCAGCATGATGCGC 960  
Db 901 TCCTTACAAGTCTTTGAGATCAGTCTGAGAAAGAGTGGCAGGATATCAGCATGATGCGC 960  
Qy 961 ATGAAGACCATTTGGGAGCAGATCCTGGCCACATCCAGACAGAGTTCGACTTCTCTTC 1020  
Db 961 ATGAAGACCATTTGGGAGCAGATCCTGGCCACATCCAGACAGAGTTCGACTTCTCTTC 1020  
Qy 1021 TGCATGGACGTGGATCAAGTCTTTCAAGACAATTTCCGGGTGGAAACTCTGGGCGCAGCTG 1080  
Db 1021 TGCATGGACGTGGATCAAGTCTTTCAAGACAATTTCCGGGTGGAAACTCTGGGCGCAGCTG 1080  
Qy 1081 GTACACAGCTCCAGGCTTGGTGTGTAAGGCCAGTCCCGAGAGATTCACCTATGAGAGG 1140  
Db 1081 GTACACAGCTCCAGGCTTGGTGTGTAAGGCCAGTCCCGAGAGATTCACCTATGAGAGG 1140  
Qy 1141 CGGGAATGTCGGCGCGGTACATTTCCATTTCGGAGAGGGGATTTTACTACCGCGCGCC 1200  
Db 1141 CGGGAATGTCGGCGCGGTACATTTCCATTTCGGAGAGGGGATTTTACTACCGCGCGCC 1200  
Qy 1201 ATTTTTCGAGGAACGCTACTACATTTCTCAACCTCACCAGGGAGTCTTTAAGGGGATC 1260  
Db 1201 ATTTTTCGAGGAACGCTACTACATTTCTCAACCTCACCAGGGAGTCTTTAAGGGGATC 1260  
Qy 1261 CTCAGGACAAAGAACATGACATAGAAGCCAGTGGCATGATGAGAGCCACCTCAACAAA 1320  
Db 1261 CTCAGGACAAAGAACATGACATAGAAGCCAGTGGCATGATGAGAGCCACCTCAACAAA 1320

Qy 1321 TACTTCTCTTTTCAACAAACCCACTAAAAATCCTATCTCCAGAGTATTGCTGGGACTATCAG 1380  
Db 1321 TACTTCTCTTTTCAACAAACCCACTAAAAATCCTATCTCCAGAGTATTGCTGGGACTATCAG 1380  
Qy 1381 ATAGCCCTGCCCTTCAGATATTAAAAGTCTCAAGGTAGCTTGGCAGACAAAAGATATAAT 1440  
Db 1381 ATAGCCCTGCCCTTCAGATATTAAAAGTCTCAAGGTAGCTTGGCAGACAAAAGATATAAT 1440  
Qy 1441 TTGGTTAGAAATAATGCTGACTTCAAAATTTGATGGAACCTTGACACTATTCTTAACCA 1500  
Db 1441 TTGGTTAGAAATAATGCTGACTTCAAAATTTGATGGAACCTTGACACTATTCTTAACCA 1500

RESULT 4  
MUSGALA  
LOCUS MUSGALA 1500 bp mRNA linear ROD 12-JUN-1993  
DEFINITION Mouse galactosyltransferase mRNA, complete cds.  
ACCESSION M26925  
VERSION M26925.1 GI:193419  
KEYWORDS galactosyl transferase.  
SOURCE Mouse F9 teratocarcinoma, cDNA to mRNA.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Larsen, R.D., Rajan, V.P., Ruff, M.M., Kukowaka-Latallo, J.,  
Cummings, R.D., and Lowe, J.B.  
Isolation of a cDNA encoding a murine UDPgalactose:beta-D-  
galactosyl-1,4-N-acetyl-D-glucosaminide alpha-1,3-  
galactosyltransferase: Expression cloning by gene transfer  
Proc. Natl. Acad. Sci. U.S.A. 86, 8227-8231 (1989)  
90046769  
COMMENT Draft entry and computer-readable copy of sequence [1] kindly  
submitted by J.B.Lowe, 08-AUG-1989.  
FEATURES  
Location/Qualifiers  
1..1500  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
277..1461  
/note="galactosyltransferase (EC 2.4.1.151)"  
/codon\_start=1  
/protein\_id="AAA37657.1"  
/db\_xref="GI:309242"  
/translation="MITMLQDLHVNKISMRSKSEKSLPSRSQSQEKINVKVKVIL  
LMLIVSTVVVFWYVNRPEVGNKQDMFSPFNKNGTHSYQEDNVGREGKRN  
GDRIEEPQLWDFNPKRPDLVTPMKQPIVWEGTYDTALKEYYATQKLTGLTVF  
AVGKYIEHYLEDFLESADMYFMVGHVIFVYMDVTSRMPVHNLPHLSLQVPIRSE  
KRWQDISMRMKTIGEHILAHIOHEVDVFLCMDVDVQFQNFVETIGQLVAOLOAAM  
YKASPEKFTYERBELSAAYTPFGEGDFYHAAJEGGTPTHILNLTRCFKGILODKKH  
DIEAQWHDHSHLNKYFLFNKPTKILSPYCWYQIGLPSDIKSVKVAWQIREYNLVRN  
NV"

BASE COUNT 406 a 349 c 374 g 371 t  
ORIGIN

Query Match 100.0%; Score 1500; DB 10; Length 1500;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTTCCCTTGTAGACTCTTCTTGGAAATCAGAAATACCGATTCTGCTGAAGACCTCGCGCT 60  
Db 1 CCTTCCCTTGTAGACTCTTCTTGGAAATCAGAAATACCGATTCTGCTGAAGACCTCGCGCT 60  
Qy 61 CTCAGGCTCTGGGAGTTTGGAAACCTGTACCTTCTTCTCTGCTGAGCCCTGCTCTCT 120  
Db 61 CTCAGGCTCTGGGAGTTTGGAAACCTGTACCTTCTTCTCTGCTGAGCCCTGCTCTCT 120  
Qy 121 AGGAGGCCAGAGCTCGACAGAACTCGGTTGCTTGTGTTTGGTGGAGGGAACACAG 180  
Db 121 AGGAGGCCAGAGCTCGACAGAACTCGGTTGCTTGTGTTTGGTGGAGGGAACACAG 180  
Qy 181 CTGACGATGAGGCTGACTTTTGAACCTCAAGACATCTGCTTACCCAGTCTCTCGAATTA 240  
Db 181 CTGACGATGAGGCTGACTTTTGAACCTCAAGACATCTGCTTACCCAGTCTCTCGAATTA 240



Db 181 CTGACGATGAGGCTGACTTTTGAACCTCAAGAGATCTGCTTACCCAGTCTCCTCGGAATTA 240

Qy 241 AGCCCTGTACTACATTTGCCGTGACCTTAAGATTTTCATGATCACTACTGCTTCAAGATCTC 300

Db 241 AGCCCTGTACTACATTTGCCGTGACCTTAAGATTTTCATGATCACTACTGCTTCAAGATCTC 300

Qy 301 CATGTCAACAAGATCTCCATGTCAAGATCCAAGTCAAGATCCAAGTCAAGATCCAAGTCAAG 360

Db 301 CATGTCAACAAGATCTCCATGTCAAGATCCAAGTCAAGATCCAAGTCAAGATCCAAGTCAAG 360

Qy 361 TCTGGATCAGACAGAGAAATAATGAATGTCAAGGGAAGTAATCTGTTGATGCTGATT 420

Db 361 TCTGGATCAGACAGAGAAATAATGAATGTCAAGGGAAGTAATCTGTTGATGCTGATT 420

Qy 421 GTCTCAACCGTGTGTCGTGTTTGGGAATATGTCAACAGAAATCCACAGGTTGCTGAG 480

Db 421 GTCTCAACCGTGTGTCGTGTTTGGGAATATGTCAACAGAAATCCACAGGTTGCTGAG 480

Qy 481 AACAGATGGCAGAGGACTGGTGGTCCCAAGCTGGTTTAAAAATGGGACCCACAGTTAT 540

Db 481 AACAGATGGCAGAGGACTGGTGGTCCCAAGCTGGTTTAAAAATGGGACCCACAGTTAT 540

Qy 541 CAAGAAGACAACCTAGAGGACGAGAGAAAGGTTAGAAATGAGATCGCATTTGAAGAG 600

Db 541 CAAGAAGACAACCTAGAGGACGAGAGAAAGGTTAGAAATGAGATCGCATTTGAAGAG 600

Qy 601 CCTCAGCTATGGGACTGGTTCAATCCAAGAACCGCCCGGATGTTTTCAGAGTGACCCCG 660

Db 601 CCTCAGCTATGGGACTGGTTCAATCCAAGAACCGCCCGGATGTTTTCAGAGTGACCCCG 660

Qy 661 TGAAGCGCCGATTTGTGGGAAGGACTTATGACACAGCTCTGCTGGAAAAAGTACTAC 720

Db 661 TGAAGCGCCGATTTGTGGGAAGGACTTATGACACAGCTCTGCTGGAAAAAGTACTAC 720

Qy 721 GCCACACAGAACTCACTGTGGGCTGACAGTGTTCCTGCTGGGAAAGTACATTTGAGCAT 780

Db 721 GCCACACAGAACTCACTGTGGGCTGACAGTGTTCCTGCTGGGAAAGTACATTTGAGCAT 780

Qy 781 TACTTGAAGACTTTCTGGAGTCTGCTGACATGCTACTTTCATGTTGGCCATCGGGTCATA 840

Db 781 TACTTGAAGACTTTCTGGAGTCTGCTGACATGCTACTTTCATGTTGGCCATCGGGTCATA 840

Qy 841 TTTTACGTATGATAGACACACCTCCGGATGCCGTGCTGACCTGAAACCCCTCTACAT 900

Db 841 TTTTACGTATGATAGACACACCTCCGGATGCCGTGCTGACCTGAAACCCCTCTACAT 900

Qy 901 TCCCTTACAAGCTTTTGAGATCAGGCTCTGAGAAGAGGTGGCAGGATATCAGCATGATGCCG 960

Db 901 TCCCTTACAAGCTTTTGAGATCAGGCTCTGAGAAGAGGTGGCAGGATATCAGCATGATGCCG 960

Qy 961 ATGAAGACCATTTGGGAGCACATCCTGGCCACATCCAGCAGGTCGACTTTCCTCTTC 1020

Db 961 ATGAAGACCATTTGGGAGCACATCCTGGCCACATCCAGCAGGTCGACTTTCCTCTTC 1020

Qy 1021 TGATGGACGTGATCAAGTCTTTCAAGACAACCTTCGGGGTGGAAACTCTGGGCCAGCTG 1080

Db 1021 TGATGGACGTGATCAAGTCTTTCAAGACAACCTTCGGGGTGGAAACTCTGGGCCAGCTG 1080

Qy 1081 GTAGCAGCTCCAGGCTGGTGGTACAAGGCCAGTCCCGAGAGTTTCACTATGAGAGG 1140

Db 1081 GTAGCAGCTCCAGGCTGGTGGTACAAGGCCAGTCCCGAGAGTTTCACTATGAGAGG 1140

Qy 1141 CGGGAACCTGTCCGCCGCTACATTTCCATTCGGAGAGGGGATTTTACTACACACGGGCC 1200

Db 1141 CGGGAACCTGTCCGCCGCTACATTTCCATTCGGAGAGGGGATTTTACTACACACGGGCC 1200

Qy 1201 ATTTTGGAGGACGGCTACTACATTTCTCAACCTCACAGGGAGTCTTTAAGGGATC 1260

Db 1201 ATTTTGGAGGACGGCTACTACATTTCTCAACCTCACAGGGAGTCTTTAAGGGATC 1260

Qy 1261 CTCAGSACAAGAAACATGACATAGAGCCAGTGCATGATCAGAGCCACCTCAACAAA 1320

Db 1261 CTCAGSACAAGAAACATGACATAGAGCCAGTGCATGATCAGAGCCACCTCAACAAA 1320

Qy 1321 TACTTCCCTTTTCAACAAACCCACTAAATCTATCTCCAGAGTATTGCTGGGACTATCAG 1380

Db 1321 TACTTCCCTTTTCAACAAACCCACTAAATCTATCTCCAGAGTATTGCTGGGACTATCAG 1380

Qy 1381 ATAGGCTGCTGCTTTTCAGATATTTAAAGTGTCAAGTGTGGCAGACAAAGAGTATAAT 1440

Db 1381 ATAGGCTGCTGCTTTTCAGATATTTAAAGTGTCAAGTGTGGCAGACAAAGAGTATAAT 1440

Qy 1441 TTGCTTTAGAAATTAATCTGCTGACTTCAAAATGTCATGGAACCTTGACACTATTCTTACCA 1500

Db 1441 TTGCTTTAGAAATTAATGTCGACTTCAAAATGTCATGGAACCTTGACACTATTCTTACCA 1500

RESULT 5

AX128259

LOCUS AX128259 3537 bp DNA linear PAT 15-MAY-2001

DEFINITION Sequence 26 from Patent WO0130992.

ACCESSION AX128259

VERSION AX128259.1 GI:14134780

KEYWORDS house mouse.

SOURCE Mus musculus

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 3537)

AUTHORS Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.

TITLE -g(a)11-3 galactosyltransferase gene and promoter

JOURNAL Patent: WO 0130992-A. 26 03-MAY-2001.

UNIV. PITTSBURGH OF THE COMMONWEALTH SYSTEM OF HIGHER EDUCATION (US)

FEATURES

source Location/Qualifiers

1..3537

/organism="Mus musculus"

/db\_xref="taxon:10090"

BASE COUNT 978 a 777 c 843 g 939 t

ORIGIN

Query Match 95.7%; Score 1434.8; DB 6; Length 3537;

Best Local Similarity 97.5%; Pred. No. 0;

Matches 1493; Conservative 0; Mismatches 2; Indels 37; Gaps 2;

Qy 1 CTTTCCCTTTGTAGACTCTTCTTGGAAATGAGAAGTACCGATTCTGCTGAAGACCTCGCGCT 60

Db 120 CTTTCCCTTTGTAGACTCTTCTTGGAAATGAGAAGTACCGATTCTGCTGAAGACCTCGCGCT 179

Qy 61 CTCAGGCTCTGGAGAGTTGGAAACCTGTACCTTCTCTCTGCTGAGAGCCCTGCCTCCTT 120

Db 180 CTCAGGCTCTGGAGAGTTGGAAACCTGTACCTTCTCTCTGCTGAGAGCCCTGCCTCCTT 239

Qy 121 AGCAGGCCAGAGCTCGACAGAA-CTCGGTTGCTTTGCTGCTTTGCTGAGGGAACACA 179

Db 240 AGCAGGCCAGAGCTCGACAGAAAGCTCGGTTGCTGCTGCTTTGCTGAGGGAACACA 299

Qy 180 GCTGACGATGAGGCTGACTTTGAACTCAAGAGATGCTTTACCCAGTCTCCTCGGAATTA 239

Db 300 GCTGACGATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCAGTCTCCTCGGAATTA 359

Qy 240 AAGCCCTGTACTACATTTGCTGACCTAAGATTTTCATGATCACTATGCTTCAAGATCT 299

Db 360 AAGCCCTGTACTACATTTGCTGACCTAAGATTTTCATGATCACTATGCTTCAAGATCT 419

Qy 300 CCATGTCAACAAGATCTCCATGTCAAGTCAAGAACCAAGTCTTCCATCCTCAAG 359

Db 420 CCATGTCAACAAGATCTCCATGTCAAGTCAAGAACCAAGTCTTCCATCCTCAAG 479

Qy 360 ATCTGGATCACAGGAGAAAAATGAATGTCAAGGAAAAAGTAATCCTGTTGATGCTGAT 419

Db 480 ATCTGGATCACAGGAGAAAAATGAATGTCAAGGAAAAAGTAATCCTGTTGATGCTGAT 539

Qy 420 TGTCTCAACCGTGGTGTGCTGTTTGGGAATATGTCAACAG----- 461

|||||

```
Db 540 TGTCTCAACCGTGGTTCGTGCTGTTTTCGGGAATATGTCAACAGCCACGACCGCTCTTCTT 599
Qy 462 -----AATTCAGAGGTGGTGTGAGAACAGATGGCAGAGGACTGGTG 503
Db 600 GTGGATATATACACAAAAATTCAGAGGTGGTGTGAGAACAGATGGCAGAGGACTGGTG 659
Qy 504 GTTCCCAAGCTGTTTAAAAATGGGACCCACAGTTTCAAGAACAGACACGTAGAACGACG 563
Db 660 GTTCCCAAGCTGTTTAAAAATGGGACCCACAGTTTCAAGAACAGACACGTAGAACGACG 719
Qy 564 GAGAGAAAGGGTAGAATCGAGATCCGATTGAAGACGCTCAGCTATGGGACTGGTTCAA 623
Db 720 GAGAGAAAGGGTAGAATCGAGATCCGATTGAAGACGCTCAGCTATGGGACTGGTTCAA 779
Qy 624 TCCAAAGAACGCCCGGATGTTTGGACAGTGACCCCGTGAAGCGCCGATGTTGTGGGA 683
Db 780 TCCAAAGAACGCCCGGATGTTTGGACAGTGACCCCGTGAAGCGCCGATGTTGTGGGA 839
Qy 684 AGGCACATTATGACACAGCTCTGCTGGAAAAAGTACTACGCCACACAGAAATCACTGTGGG 743
Db 840 AGGCACATTATGACACAGCTCTGCTGGAAAAAGTACTACGCCACACAGAAATCACTGTGGG 899
Qy 744 GCTGACAGTGTCTGCTGGAAAGTACATTGAGCATTACTTACAGACATTTCTGGAGTC 803
Db 900 GCTGACAGTGTCTGCTGGAAAGTACATTGAGCATTACTTACAGACATTTCTGGAGTC 959
Qy 804 TGCTGACATGTACTTTCATGTTGGCCATCGGGTCATATTTACGTCATGATAGACGAC 863
Db 960 TGCTGACATGTACTTTCATGTTGGCCATCGGGTCATATTTACGTCATGATAGACGAC 1019
Qy 864 CTCCCGGATGCTGCTGACACCTGAACCTCTACATTCCTTACAACTCTTTGAGATCAG 923
Db 1020 CTCCCGGATGCTGCTGACACCTGAACCTCTACATTCCTTACAACTCTTTGAGATCAG 1079
Qy 924 GTCTGAGAAAGGTGGCAGGATATCAGCATGATCGCATGAAGACCATTTGGGAGCACAT 983
Db 1080 GTCTGAGAAAGGTGGCAGGATATCAGCATGATCGCATGAAGACCATTTGGGAGCACAT 1139
Qy 984 CTTGGCCACATCCAGCAGGAGTGCAGTTCCTCTTCTGCTGATGGACGCTGGATCAAGTCTT 1043
Db 1140 CTTGGCCACATCCAGCAGGAGTGCAGTTCCTCTTCTGCTGATGGACGCTGGATCAAGTCTT 1199
Qy 1044 TCAAGACAATCTCGGGGTGAAACTCTGGGGCAGCTGGTAGCACAGCTCCAGGCGCTGGTG 1103
Db 1200 TCAAGACAATCTCGGGGTGAAACTCTGGGGCAGCTGGTAGCACAGCTCCAGGCGCTGGTG 1259
Qy 1104 GTACAAGCCAGTCCCGAGAGTTACCTATATGAGAGCGGGAACCTGTCGGCGCGGTACAT 1163
Db 1260 GTACAAGCCAGTCCCGAGAGTTACCTATATGAGAGCGGGAACCTGTCGGCGCGGTACAT 1319
Qy 1164 TCCATTCCGAGAGGGGATTTTACTACCGCGGCCATTTTGGAGGAAACGCGCTACTCA 1223
Db 1320 TCCATTCCGAGAGGGGATTTTACTACCGCGGCCATTTTGGAGGAAACGCGCTACTCA 1379
Qy 1224 CATTTCTCAACCTCACAGGAGGTGCTTTAAGGGGATCCTCCAGGACAAGAAACATGACAT 1283
Db 1380 CATTTCTCAACCTCACAGGAGGTGCTTTAAGGGGATCCTCCAGGACAAGAAACATGACAT 1439
Qy 1284 AGAAGCCAGTGGCATGATGAGAGCCACCTCAACAAATACTTCTTTTCAACAACCCAC 1343
Db 1440 AGAAGCCAGTGGCATGATGAGAGCCACCTCAACAAATACTTCTTTTCAACAACCCAC 1499
Qy 1344 TAAATCCTATCTCCAGAGATTTGCTGGGACTATCAGATAGGCGCTGCTTTCAGATATTA 1403
Db 1500 TAAATCCTATCTCCAGAGATTTGCTGGGACTATCAGATAGGCGCTGCTTTCAGATATTA 1559
Qy 1404 AAGTGTCAAGGTAGCTTTGGCAGACAAAAGAGTATAATTTGGTTAGAAATATGCTGACT 1463
Db 1560 AAGTGTCAAGGTAGCTTTGGCAGACAAAAGAGTATAATTTGGTTAGAAATATGCTGACT 1619
Qy 1464 TCAAAATTTGATGGAAACTTTGACACTATTTCT 1495
Db 1620 TCAAAATTTGATGGAAACTTTGACACTATTTACT 1651
```

## RESULT 6

```
AF297615 LOCUS 3537 bp mRNA linear ROD 05-SEP-2001
DEFINITION Mus musculus alpha-1,3-galactosyltransferase gene, complete cds.
ACCESSION AF297615
VERSION AF297615.1 GI:15419873
KEYWORDS house mouse.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3537)
AUTHORS Koike,C., Friday,R., Fung,J.J., Starzl,T.E. and Trucco,M.
TITLE Comparison of the regulatory regions of the
alpha1,3galactosyltransferase gene between murine and porcine
species
JOURNAL Transplant. Proc. 33 (1-2), 710-711 (2001)
MEDLINE 21168575
PUBMED 11267031
REFERENCE 2 (bases 1 to 3537)
AUTHORS Koike,C., Starzl,T.E. and Trucco,M.
TITLE Isolation and characterization of the regulatory region of the
murine alpha-1,3-galactosyltransferase gene
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 3537)
AUTHORS Koike,C., Starzl,T.E. and Trucco,M.
TITLE Direct Submission
JOURNAL Submitted (21-AUG-2000) Surgery, University of Pittsburgh, 3601
Fifth Ave., Pittsburgh, PA 15215, USA
FEATURES
source
1..3537
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
502..1617
/codon_start=1
/product="alpha-1,3-galactosyltransferase"
/protein_id="AAK97206.1"
/db_xref="GI:15419874"
/translation="MNVGKVIILLIIVTVVVVFVEYVNSPDGSELIYHTIPEVG
ENRWKDMFPSPSWFKNGTHSYQEDNVGREGKRNDRLEEPQLWDFNFKNRPDLVT
VTPKQIWEVGTIDTALKEYYATOKLTGLTVFVAVGYKIEHYLEDFLDSADMYFMV
GHRVIFYMIDDTSRMPVHNLPLHSLQVFEIRSEKRWODISMHRKTIGEHTLAHIQ
HEYDFLCMDVDQVFNFGVETLGVLQALQAWKYKASPEKTYRRELISAAYIPFG
EGDFYVHAATFGGTPTHLNLTRECFKGLQDKKHDEAQMHDESHLNKYLFLNKPTK
ILSPEYCDYQIGLPSDIKSVKVMQTKYNLVRNV"
polyA_signal 3419..3424
polyA_site 3521..3527
BASE COUNT 978 a 777 c 843 g 939 t
ORIGIN
```

```
Query Match 95.7%; Score 1434.8; DB 10; Length 3537;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 1493; Conservative 0; Mismatches 2; Indels 37; Gaps 2;

Qy 1 CCTTCCCTCTAGACTCTTCTTGAATGAGAAGTACCGATTCTGCTGAAGACCTCGCGCT 60
Db 120 CCTTCCCTCTAGACTCTTCTTGAATGAGAAGTACCGATTCTGCTGAAGACCTCGCGCT 179
Qy 61 CTCAGGCTCTGGGAGTTGGAAACCTGTACCTTCCTTCTCTGCTGAGCCCTGCCTCTT 120
Db 180 CTCAGGCTCTGGGAGTTGGAAACCTGTACCTTCCTTCTCTGCTGAGCCCTGCCTT 239
Qy 121 AGGAGCCGACAGCTCGACAGAA-CTCGGTTGCTTTGCTGTTGCTTTGGAGGAAACACA 179
Db 240 AGGAGCCGACAGCTCGACAGAACTCGGTTGCTTTGCTGTTGCTTTGGAGGAAACACA 299
Qy 180 GCTCAGCATGAGGCTGACTTTGAACCTCAAGAGATCTGCTTACCACGATCTCCTCGAATTA 239
```

Db 300 GCTGACGATGAGGCTGACTTTTGAACCTCAAGAGATCTGCTTTACCCCGAGTCTCCTGGGAATTA 359  
QY 240 AAGGCGCTGTACTACATTTTGCCTGGACCTAAGATTTTTCATGATCACTATGCTTCAAGATCT 299  
Db 360 AAGGCGCTGTACTACATTTTGCCTGGACCTAAGATTTTTCATGATCACTATGCTTCAAGATCT 419  
QY 300 CCATGTCAACAAGATCTCCATGTCAAGATCAAGTCCAGAACAAAGTCTTCCATCCTCAAG 359  
Db 420 CCATGTCAACAAGATCTCCATGTCAAGATCAAGTCCAGAACAAAGTCTTCCATCCTCAAG 479  
QY 360 ATCTGGATCACAGAGAAATAATGAATGTCAAGGGAAAGATTAATCTGTGTGATGCTGAT 419  
Db 480 ATCTGGATCACAGAGAAATAATGAATGTCAAGGGAAAGATTAATCTGTGTGATGCTGAT 539  
QY 420 TGTCTCAACCGTGGTGTCTGCTGTTTTGGGAATATGTCAACAG----- 461  
Db 540 TGTCTCAACCGTGGTGTCTGCTGTTTTGGGAATATGTCAACAGCCAGACGGCTCTTCTT 599  
QY 462 -----AATTCAGAGTGTGTTGAGACAGATGGCAGAAAGACTGGTG 503  
Db 600 GTGGATATATCACACAAAATTCAGAGGTTGGTGAGAACAGATGGCAGAAAGACTGGTG 659  
QY 504 GTTCCCAAGCTGGTTTAAAAATGGACCCACACTTATCAAGAGACAAACGTCAAGAGACG 563  
Db 660 GTTCCCAAGCTGGTTTAAAAATGGACCCACAGTTATCAAGAGACAAACGTCAAGAGACG 719  
QY 564 GAGAGAAAAGGTAGAAATGGAGATGGCATTTGAAGAGCCCTCAGCTATGGGACTGGTTCAA 623  
Db 720 GAGAGAAAAGGTAGAAATGGAGATGGCATTTGAAGAGCCCTCAGCTATGGGACTGGTTCAA 779  
QY 624 TCCAAAGAACCGCCGGATGTTTTGACAGTGCACCCCGTGAAGCGCCGATTTGTGGGA 683  
Db 780 TCCAAAGAACCGCCGGATGTTTTGACAGTGCACCCCGTGAAGCGCCGATTTGTGGGA 839  
QY 684 AGGCACTTATGACACAGCTCTGCTGGAAGAGTACTAGCCACACAGAACTCACTGTGG 743  
Db 840 AGGCACTTATGACACAGCTCTGCTGGAAGAGTACTAGCCACACAGAACTCACTGTGG 899  
QY 744 GCTGACAGTGTTCCTGCGGAAAGTACATTTGAGCATTACTTGAAGACTTTCTGGAGTC 803  
Db 900 GCTGACAGTGTTCCTGCGGAAAGTACATTTGAGCATTACTTGAAGACTTTCTGGAGTC 959  
QY 804 TGCTGACATGACTTTCATGTTGGCATCGGGTCAATTTTACGTCATGATGACGACAC 863  
Db 960 TGCTGACATGACTTTCATGTTGGCATCGGGTCAATTTTACGTCATGATGACGACAC 1019  
QY 864 CTCCGGATGCTGCTGACACCTGAACCCCTCTACATTCCTTACAAGTCTTTGAGATCAG 923  
Db 1020 CTCCGGATGCTGCTGACACCTGAACCCCTCTACATTCCTTACAAGTCTTTGAGATCAG 1079  
QY 924 GTCTGAGAGAGGTGGCAGGATATCAGCATGATCGCATGAAGACCATTTGGGGAGGACAT 983  
Db 1080 GTCTGAGAGAGGTGGCAGGATATCAGCATGATCGCATGAAGACCATTTGGGGAGGACAT 1139  
QY 984 CTGTGGCCACATCCAGACAGGTGCTGCTTCTTCTGATGGAGTGGATCAAGTCTTT 1043  
Db 1140 CTGTGGCCACATCCAGACAGGTGCTGCTTCTTCTGATGGAGTGGATCAAGTCTTT 1199  
QY 1044 TCAAGACAACCTTCGGGTGAAAACCTCTGGCCAGCTGGTAGCAGCTCCAGGCTGGTG 1103  
Db 1200 TCAAGACAACCTTCGGGTGAAAACCTCTGGCCAGCTGGTAGCAGCTCCAGGCTGGTG 1259  
QY 1104 GTACAGGCGCAGTCCCGAGAAAGTTACCTATGAGAGCGGGAACCTGTGCGCCCGCTACAT 1163  
Db 1260 GTACAGGCGCAGTCCCGAGAAAGTTACCTATGAGAGCGGGAACCTGTGCGCCCGCTACAT 1319  
QY 1164 TCCATTCGGAGAGGGGATTTTACTACACCGCGGCCATTTTGGAGGAACGCTTACTCA 1223  
Db 1320 TCCATTCGGAGAGGGGATTTTACTACACCGCGGCCATTTTGGAGGAACGCTTACTCA 1379  
QY 1224 CATTCACAACCTCACAGGGAGTGCCTTTAAGGGATCTCCAGGACAGAAACATGACAT 1283  
Db 1380 CATTCACAACCTCACAGGGAGTGCCTTTAAGGGATCTCCAGGACAGAAACATGACAT 1439

QY 1284 AGAAGCCAGTGGCATGATGAGAGCCACCTCAACAATACTTCTCTTTTCAACAACCCAC 1343  
Db 1440 AGAAGCCAGTGGCATGATGAGAGCCACCTCAACAATACTTCTCTTTTCAACAACCCAC 1499  
QY 1344 TAAATCTCTATCTCCAGAGTATTGCTGGGACTATCAGATAGGCTGCTTCAGATATTA 1403  
Db 1500 TAAATCTCTATCTCCAGAGTATTGCTGGGACTATCAGATAGGCTGCTTCAGATATTA 1559  
QY 1404 AAGTGTCAAGGTAGCTTGGCAGACAAAGAGTATAATTTGTTAGAAATAATGCTGACT 1463  
Db 1560 AAGTGTCAAGGTAGCTTGGCAGACAAAGAGTATAATTTGTTAGAAATAATGCTGACT 1619  
QY 1464 TCAATTTGTGATGGAACCTTGACACTATTCT 1495  
Db 1620 TCAATTTGTGATGGAACCTTGACACTATTACT 1651  
RESULT 7  
MUSGLYTNs  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
FEATURES  
Source  
1. .3450  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/cell\_line="C127"  
1. .3450  
/gene="Gata-1"  
445. .1560  
/gene="Gata-1"  
/note="positions 463-495 code for the transmembrane domain"  
/codon\_start=1  
/product="alpha-1,3-galactosyltransferase"  
/protein\_id="AA37711.1"  
/db\_xref="GI:457142"  
/translation="MNKCKVILLMLIVSTVVVFWYVNSPDGSLWLYHTKIPVG  
ENRWKDNWFPSPKNGTHSYOEDNVEGRKRCNDRIEEPQLMDWPNKRPDVL  
VTPWKAPVWEGTYDTALLKTYATOKLTVLTAVGKTYIEHYLEDPLSDADMYFW  
GHRVIFVYMDIDTSRPMVHLNPLSLQVFEINSEKQWDISMMRKTIGEHLAHIQ  
HEVDFLCMDVDQVDFNVEFTLGQVLAQAWMYKASPERFTYRRLSAAYIPFG  
EDDFVYHAAIFGFTPTTHILNLTRECFKGLDOKHIDIEAQWHDHSLNLYFLFNKPTK  
ILSPEYCDWYQIGLPSIDIKSVKAWOTKEYNLVRNV"  
BASE COUNT 954 a 762 c 826 g 908 t  
ORIGIN chromosome 2 (centromeric region).  
Query Match 94.6%; Score 1419.6; DB 10; Length 3450;  
Best Local Similarity 97.3%; Pred. No. 0;  
Matches 1490; Conservative 0; Mismatches 4; Indels 38; Gaps 3;  
QY 1 CCTTCCCTTGTAGACTCTTCTTGAATGAGAAGTACCGATTCTCTGCTGAAGACCTCGCGCT 60  
Db 64 CCTTCCCTTGTAGACTCTTCTTGAATGAGAAGTACCGATTCTCTGCTGAAGACCTCGCGCT 123  
QY 61 CTGAGGCTCGGAGTGGAAACCTGTACCTTCTCTTCTGCTGAGCCCTGCTCCTT 120  
Db 124 CTCAGGCTCGGAGTGGAAACCTGTACCTTCTCTTCTGCTGAGCCCTGCTCCTT 183

```
Qy 121 AGCAGGCCAGAGCTCGACAGAA-CTCGGTTGCTTTGCTGTTTCTGTTGGAGGGAACACA 179
Db 184 CGCAGGCCAGAGCTCGACAGAACTCGGTTGCTTTGCTGTTTCTGTTGGAGGGAACACA 243
Qy 180 GCTGACGATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCAGTCTCCTCGAATTA 239
Db 244 GCTGACGATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCAGTCTCCTCGAATTA 303
Qy 240 AAGGCTGTACTACATTTGCTGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCT 299
Db 304 AAGGCTGTACTAC-CTTGCTGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCT 362
Qy 300 CCATGTCAACAAGATCTCCATGTCAAGATCCAAGTCAAGAAAGTCTTCCATCCTCAAG 359
Db 363 CCATGTCAACAAGATCTCCATGTCAAGATCCAAGTCAAGAAAGTCTTCCATCCTCAAG 422
Qy 360 ATCTGGATCACAGAGAAAATAATGAATGTCAAGGGAAGTAATCTCTGTGATGCTGAT 419
Db 423 ATCTGGATCACAGAGAAAATAATGAATGTCAAGGGAAGTAATCTCTGTGATGCTGAT 482
Qy 420 TGTCTCAACCGTGTGCTGTTTGGGAATATGTCAACAG----- 461
Db 483 TGTCTCAACCGTGTGCTGTTTGGGAATATGTCAACAGCCGACGCTCTTCTT 542
Qy 462 -----AATTCAGAGTTGGTGAACAACAGATGGCAGAGGACTGGTG 503
Db 543 GTGCATATATCACACAAAAATTCAGAGTTGGTGAACAACAGATGGCAGAGGACTGGTG 602
Qy 504 GTTCCCAAGCTGTTTAAAAATGGGACCCACAGTTATCAAGAACAACAGCTAGAAGACG 563
Db 603 GTTCCCAAGCTGTTTAAAAATGGGACCCACAGTTATCAAGAACAACAGCTAGAAGACG 662
Qy 564 GAGAAAGGTTAGAAATGAGATCGCATTGAAGAGCTCAGCTATGGGACTGGTCAA 623
Db 663 GAGAAAGGTTAGAAATGAGATCGCATTGAAGAGCTCAGCTATGGGACTGGTCAA 722
Qy 624 TCCAAGAAGCCCGGATGTTTGAAGTGAACCCGTTGAAGGCGCGGATGTTGTGGGA 683
Db 723 TCCAAGAAGCCCGGATGTTTGAAGTGAACCCGTTGAAGGCGCGGATGTTGTGGGA 782
Qy 684 AGGCACATTATGACACAGCTCTGCTGGAAAGTACTAGCCACACAGAACTCACTGTGG 743
Db 783 AGGCACATTATGACACAGCTCTGCTGGAAAGTACTAGCCACACAGAACTCACTGTGG 842
Qy 744 GCTGACAGTGTGCTGTGGAAAGTACATTGAGCATTACTTAGAAGACTTCTTGGAGTC 803
Db 843 GCTGACAGTGTGCTGTGGAAAGTACATTGAGCATTACTTAGAAGACTTCTTGGAGTC 902
Qy 804 TGCTGACATGTACTTTCATGTTGGCCATCGGGTCAATATTTACGTCATGATAGACGAC 863
Db 903 TGCTGACATGTACTTTCATGTTGGCCATCGGGTCAATATTTACGTCATGATAGACGAC 962
Qy 864 CTCCCGGATGCTCTGCTGACCTGAACCTCTACATTCTTACAAGTCTTTGAGATCAG 923
Db 963 CTCCCGGATGCTCTGCTGACCTGAACCTCTACATTCTTACAAGTCTTTGAGATCAG 1022
Qy 924 GTCTGAGAAGAGTGGGAGGATATCAGCATGATCGCATGAAGACCATTTGGGGAGCACAT 983
Db 1023 GTCTGAGAAGAGTGGGAGGATATCAGCATGATCGCATGAAGACCATTTGGGGAGCACAT 1082
Qy 984 CTTGGCCACATCCAGACAGAGTTCGACTTCCTCTTCTGCAATGACGCTGGAATCAAGTCT 1043
Db 1083 CTTGGCCACATCCAGACAGAGTTCGACTTCCTCTTCTGCAATGACGCTGGAATCAAGTCT 1142
Qy 1044 TCAAGACAACCTTCGGGTTGGAACCTCTGGCCAGCTGTAAGCAGCTTCCAGGCTGGTG 1103
Db 1143 TCAAGACAACCTTCGGGTTGGAACCTCTGGCCAGCTGTAAGCAGCTTCCAGGCTGGTG 1202
Qy 1104 GTACAAGCCAGTCCCGAGAGTTTCACTATGAGAGCGGGAACCTGTGGCGCGGTACAT 1163
Db 1203 GTACAAGCCAGTCCCGAGAGTTTCACTATGAGAGCGGGAACCTGTGGCGCGGTACAT 1262
```

```
Qy 1164 TCCATTCGGAGAGGGGATTTTTACTACCGCGGCCATTTTTTGGAGAACGCGCTACTCA 1223
Db 1263 TCCATTCGGAGAGGGGATTTTTACTACCGCGGCCATTTTTTGGAGAACGCGCTACTCA 1322
Qy 1224 CATTTCTCAACCTCACAGGAGTCTTTTAAGGGGATCTCCAGGACAAGAATGACAT 1283
Db 1323 CATTTCTCAACCTCACAGGAGTCTTTTAAGGGGATCTCCAGGACAAGAATGACAT 1382
Qy 1384 AGAAGCCAGTGGCATGATGAGGCCACTCAACAATACTTCCCTTTCAACAACCCAC 1343
Db 1383 AGAAGCCAGTGGCATGATGAGGCCACTCAACAATACTTCCCTTTCAACAACCCAC 1442
Qy 1344 TAAATCTCTATCTCCAGAGTATGCTGGGACTATCAGATAGGCGCTTCAGATATTTAA 1403
Db 1443 TAAATCTCTATCTCCAGAGTATGCTGGGACTATCAGATAGGCGCTTCAGATATTTAA 1502
Qy 1404 AAGTGTCAAGTACCTTGGCAGACAAAAGAGTATAATTTGGTTAGAAATAATGCTGACT 1463
Db 1503 AAGTGTCAAGTACCTTGGCAGACAAAAGAGTATAATTTGGTTAGAAATAATGCTGACT 1562
Qy 1464 TCAAAATTGTGATGGAACCTTGACACATATTTCT 1495
Db 1563 TCAAAATTGTGATGGAACCTTGACACATATTTCT 1594

RESULT 8
AR066335
LOCUS AR066335 3450 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 9 from patent US 5849991.
ACCESSION AR066335
VERSION AR066335.1 GI:5996551
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3450)
AUTHORS d'Aplice,A.J.F., Pearse,M.J., Robins,A.J., Crawford,R.J. and
Rathjen,P.D.
TITLE Mice homozygous for an inactivated .alpha. 1,3-galactosyl
transferase gene
JOURNAL Patent: US 5849991-A 9 15-DEC-1998;
FEATURES Location/Qualifiers
source 1. 3450
BASE COUNT 954 a 762 c 826 g 908 t
ORIGIN

Query Match 93.9%; Score 1408.4; DB 6; Length 3450;
Best Local Similarity 96.8%; Pred. No. 0;
Matches 1483; Conservative 0; Mismatches 11; Indels 38; Gaps 3;

Qy 1 CTTTCCCTTTGTAGACTCTTCTTGAATGAGAAGTACCGATTCTGCTGAAGACCTCGCGCT 60
Db 64 CTTTCCCTTTGTAGACTCTTCTTGAATGAGAAGTACCGATTCTGCTGAAGACCTCGCGCT 123
Qy 61 CTCAGGCTCTGGGAGTTGGAAACCTTGTACCTTCTCTTCTCTGCTGAGCCCTGCCCTCT 119
Db 124 CTCAGGCTCTGGGAGTTGGAAACCTTGTACCTTCTCTCTGCTGAGCCCTGCCCTCT 183
Qy 120 TAGCGAGCCAGAGCTCGACAGAACTCGGTTGCTTTGCTGTTTGGAGGGAACACA 179
Db 184 TCGCAGGCCAGAGCTGACAGAGCTCGGTTGCTTTGCTGTTTGGAGGGAACACA 243
Qy 180 GCTCAGCATGAGGCTGACTTTTGAACCTCAAGAGATCTGCTTACCCAGTCTCCTCGAATTA 239
Db 244 GCTCAGCATGAGGCTGACTTTTGAACCTCAAGAGATCTGCTTACCCAGTCTCCTCGAATTA 303
Qy 240 AAGGCTGTACTACATTTTGGCTGGACCTTAAGATTTTTCATGATCACTATGCTTCAAGATCT 299
Db 304 AAGGCTGTACTAC-CTTGGCTGGACCTTAAGATTTTTCATGATCACTATGCTTCAAGATCT 362
Qy 300 CCATGTCAACAAGATCTCCATGTCAAGTCCAAGTCAAGAAAGTCTTCCATCCTCAAG 359
```

```
|||||
363 CCATGTCAACAGATCTCCATGTCAAGATCCAAAGTCAGAAACAGCTCTCCATCTCCACG 422
|||||
360 ATCTGGATCACAGAGAAATAATGAATGTCAAGGGAAGATTAATCTGTGTGATCTGAT 419
|||||
423 ATCTGGATCACAGAGAAATAATGAATGTCAAGGGAAGATTAATCTGTGTGATCTGAT 482
|||||
420 TGCTCTCAACCGTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 461
|||||
483 TGCTCTCAACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 542
|||||
462 -----AATTCACAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 503
|||||
543 GTGGATATATCACACAAAAATCCAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 602
|||||
504 GTTCCCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 563
|||||
603 GTTCCCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 662
|||||
564 GAGAGAAAAGGTAGAAATGAGATCGCATTTGAAGAGCCCTCAGCTATGGGACTGGTCAA 623
|||||
663 GAGAGAAAAGGTAGAAATGAGATCGCATTTGAAGAGCCCTCAGCTATGGGACTGGTCAA 722
|||||
624 TCCAAAGAACCGCCCGGATGTTTGACAGTGACAGTACCCCGTGAAGCGCGGATTTGTGGGA 683
|||||
723 TCCAAAGAACCGCCCGGATGTTTGACAGTGACAGTACCCCGTGAAGCGCGGATTTGTGGGA 782
|||||
684 AGGCACCTTATGACACAGCTCTGCTGTGGAAGTACTACGCCACACAGAACTCACTGTGGG 743
|||||
783 AGGCACCTTATGACACAGCTCTGCTGTGGAAGTACTACGCCACACAGAACTCACTGTGGG 842
|||||
744 GCTGACAGTGTGCTGTGGGAAGTACATTTAGACATTTACTTGAAGACTTTCTGGAGTC 803
|||||
843 GCTGACAGTGTGCTGTGGGAAGTACATTTAGACATTTACTTGAAGACTTTCTGGAGTC 902
|||||
804 TGCTGACATGATCTCATGTTGGCCATCGGGTCATATTTTAGTCATGATGACACACAC 863
|||||
903 TGCTGACATGATCTCATGTTGGCCATCGGGTCATATTTTAGTCATGATGACACAC 962
|||||
864 CTCGGGATGCTGTGTCACCTGAAACCTCTACATTTCTTCAAGTCTTTGAGATCAG 923
|||||
963 CTCGGGATGCTGTGTCACCTGAAACCTCTACATTTCTTCAAGTCTTTGAGATCAG 1022
|||||
924 GTCTGAGAAGAGTGGCAGGATATCAGCATGATGCGCATGAAGACCATTTGGGAGCACAT 983
|||||
1023 GTCTGAGAAGAGTGGCAGGATATCAGCATGATGCGCATGAAGACCATTTGGGAGCACAT 1082
|||||
984 CCTGGCCACATCCAGCACAGTCTGCTCTTCTGCTGATGGACGTGATCAAGTCTTT 1043
|||||
1083 CCTGGCCACATCCAGCACAGTCTGCTCTTCTGCTGATGGACGTGATCAAGTCTTT 1142
|||||
1044 TCAAGACAACCTTCGGGGTGAAATCTCTGGGCCAGCTGCTAGCACAGCTCCAGCGCTGGTG 1103
|||||
1143 TCAAGACAACCTTCGGGGTGAAATCTCTGGGCCAGCTGCTAGCACAGCTCCAGCGCTGGTG 1202
|||||
1104 GTACAAGGCCAGTCCCAGAGAGTTCACTATGAGAGCGGGAACTGTGCGCCCGCTACAT 1163
|||||
1203 GTACAAGGCCAGTCCCAGAGAGTTCACTATGAGAGCGGGAACTGTGCGCCCGCTACAT 1262
|||||
1164 TCCATTCCGAGAGGGGGATTTTACTACACAGCGGGCCATTTTTCGAGGAACGCTACTCA 1223
|||||
1263 TCCATTCCGAGAGGGGGATTTTACTACACAGCGGGCCATTTTTCGAGGAACGCTACTCA 1322
|||||
1224 CATTTCTCAACTTCACAGGGAGTCTTTAAGGGATCTCTCAGGACAAAGAAACATGACAT 1283
|||||
1323 CATTTCTCAACTTCACAGGGAGTCTTTAAGGGATCTCTCAGGACAAAGAAACATGACAT 1382
|||||
1284 AGAAGCCCACTGGCATGATGAGAGCCACCTCAACAATACTTCTTTTCAACAACCCAC 1343
|||||
1383 AGAAGCCCACTGGCATGATGAGAGCCACCTCAACAATACTTCTTTTCAACAACCCAC 1442
|||||
1344 TAAATCTCTATCTCCAGAGTATTTGCTGGGACTATCAGATAGGCGTCTTCAGATATTAA 1403
|||||
```

```
Db 1443 TAAATCTCTATCTCCAGAGTATTTGCTGGGACTATCAGATAGGCGCTCCCTTCAGATATTAA 1502
QY 1404 AACTGTCAAGGTAGCTTGGCAGACAAAGAGTATAATTTGTTTGTAGAAATTAATCTCTGACT 1463
|||||
Db 1503 AACTGTCAAGGTAGCTTGGCAGACAAAGAGTATAATTTGTTTGTAGAAATTAATCTCTGACT 1562
|||||
QY 1464 TCAATTTGTATGGAACTTTGACACTATTTCCT 1495
|||||
Db 1563 TCAATTTGTATGGAACTTTGACACTATTACT 1594
|||||

RESULT 9
BC006810 3315 bp mRNA linear ROD 12-JUL-2001
LOCUS Mus musculus, clone MGC:11545 IMAGE:3154188, mRNA, complete cds.
DEFINITION BC006810
ACCESSION BC006810.1 GI:13905053
VERSION MGC.
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3315)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (27-APR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: villalona@bcm.tmc.edu.
Villalona, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gibbs,R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 16 Row: c Column: 15
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 193419.
FEATURES
Location/Qualifiers
1..3315
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="MGC:11545 IMAGE:3154188"
/tissue_type="Mammary tumor. Brca1-/-fl; MMTV-Cre model. 10
months old, gross tissue."
/clone_lib="NCI_CGAP_Mam3"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
252..1370
/codon_start=1
/product="Unknown (protein for MGC:11545)"
/protein_id="AAH06810.1"
/db_xref="GI:13905054"
/translaton="MITMLQDLHVKNKISMSRSKSETSLPSSRSOSOEKIMNVKGVIL
LMIVSTVVVFWEYVNRTHSTQEDNVEGRNRKGRNGDRTEEPQLMDFNPKNRPDVL
TYPWKAPIVWEGTYDTALLEKYATOKLTGLTVFAVGKYIEHYLEDFLESADMYFM
VGRVIFYVYIDDTSRMPVHLNPLSLQVFEIRSEKRMQDISMRMKTIGEHILAH
QHEVDLFQCMVDQVDFNFGVETLGQLVAQLQAMWKASPEKFTYERRLESAYTIPF
GEGDFVYHAAITGGTPTHTLNLTRFCKTGKTLQDKKHDIEAQWHDIESILNKYFLFNKPT
KILSPCYCHDYQIGLPSDIKSVKVAQTKREYNLVRNV"

BASE COUNT 937 a 719 c 782 g 877 t
ORIGIN
```

Query Match		88.4%	Score 1326.4;	DB 10;	Length 3315;
Best Local Similarity		95.48;	Pred. No. 0;		
Matches 1403; Conservative		0;	Mismatches	1;	Indels 66; Gaps 1;
Qy	26	ATGAGAAGTACCGATCTCTGCTGAAGACCTCGCGCTCTCAGGCTCTGGAGTTGGAACCCCT	85		
Db	1	ATGAGAAGTACCGATCTCTGCTGAAGACCTCGCGCTCTCAGGCTCTGGAGTTGGAACCCCT	60		
Qy	86	GTACCTCTCTCTCTGCTGAGCCCTGCTCTTACGAGCCGAGAGCTCGACAGAACT	145		
Db	61	GTACCTCTCTCTCTGCTGAGCCCTGCTCTTACGAGCCGAGAGCTCGACAGAACT	120		
Qy	146	CGGTTCCTCTCTCTGCTTGGAGGGAACACAGCTGACGATGAGGCTGACTTTGAACT	205		
Db	121	CGGTTCCTCTCTCTGCTTGGAGGGAACACAGCTGACGATGAGGCTGACTTTGAACT	180		
Qy	206	CAAGAGATCTCTTACCCAGCTCTCTGGAATTAAGGSCCTGTACTACATTTGCTTGGAC	265		
Db	181	CAAGAGATCTCTTACCCAGCTCTCTGGAATTAAGGSCCTGTACTACATTTGCTTGGAC	240		
Qy	266	CTAAGATTTTCAATGATCACTATGCTTCAGATCTCCATGTCACAAAGATCTCCATGTCAA	325		
Db	241	CTAAGATTTTCAATGATCACTATGCTTCAGATCTCCATGTCACAAAGATCTCCATGTCAA	300		
Qy	326	GATCCAGTCAAGAACAGTCTTCCATCTCAAGATCTGGATCAAGGAGAAATATGA	385		
Db	301	GATCCAGTCAAGAACAGTCTTCCATCTCAAGATCTGGATCAAGGAGAAATATGA	360		
Qy	386	ATGTCAAGGGAAAAAGTAACTCTGTGTGATGCTGATGCTCAACCGTGGTTCGTGTTTT	445		
Db	361	ATGTCAAGGGAAAAAGTAACTCTGTGTGATGCTGATGCTCAACCGTGGTTCGTGTTTT	420		
Qy	446	GGGAATATGTCACAGAAATTCAGAGGTTGGTGAGACAGATGGCAGAGAACTGGTGGT	505		
Db	421	GGGAATATGTCACAA-----	435		
Qy	506	TCCCAAGCTGGTTTAAAAATGGGACCCACAGTTATCAAGAGACAACTAGAGGACGGA	565		
Db	436	-----GGACCCACAGTTATCAAGAGACAACTAGAGGACGGA	474		
Qy	566	GAGAAAAGGTAGAAATGGAGATCGCATTTGAAGAGCCCTCAGCTATGGGACCTGGTTCAATC	625		
Db	475	GAGAAAAGGTAGAAATGGAGATCGCATTTGAAGAGCCCTCAGCTATGGGACCTGGTTCAATC	534		
Qy	626	CAAGAACCCGCCGATGTTTGACAGTGACCCGCTGGAAGCCCGGATTCGTGGGAAG	685		
Db	535	CAAGAACCCGCCGATGTTTGACAGTGACCCGCTGGAAGCCCGGATTCGTGGGAAG	594		
Qy	686	GCACCTATGACACAGCTCTGCTGGAAAAGTACTACGCCACACAGAACTCACTGTGGGCG	745		
Db	595	GCACCTATGACACAGCTCTGCTGGAAAAGTACTACGCCACACAGAACTCACTGTGGGCG	654		
Qy	746	TGACAGTGTTCCTGTGGAAAGTACATTTAGACATTTACTTTAGAGACTTTCTGGAGTCTG	805		
Db	655	TGACAGTGTTCCTGTGGAAAGTACATTTAGACATTTACTTTAGAGACTTTCTGGAGTCTG	714		
Qy	806	CTGACATGACTTCTATGCTTGGCCATCGGTCATATTTTACGTCTATGATAGACACACCT	865		
Db	715	CTGACATGACTTCTATGCTTGGCCATCGGTCATATTTTACGTCTATGATAGACACACCT	774		
Qy	866	CCCGATGCCCTGCTGTCACCTGAAACCCCTCTACATTCCTTACAAGCTTTTGGAGATCAGGT	925		
Db	775	CCCGATGCCCTGCTGTCACCTGAAACCCCTCTACATTCCTTACAAGCTTTTGGAGATCAGGT	834		
Qy	926	CTGAGAAGAGGTGGCAGGATATCAGCATGATGCGCATGAAGACCATTTGGGAGACATCC	985		
Db	835	CTGAGAAGAGGTGGCAGGATATCAGCATGATGCGCATGAAGACCATTTGGGAGACATCC	894		
Qy	986	TGGCCCATCCAGCACAGCTCGACTTCTCTCTGATGAGCTGGATCAAGTCTTTC	1045		
.Db	895	TGGCCCATCCAGCACAGCTCGACTTCTCTCTGATGAGCTGGATCAAGTCTTTC	954		

Qy	1046	AAGAAACTTCGGGTGGAAACTCTGGGCCAGCTGGTAGCAGCTCCAGCCCTGGTGGT	1105		
Db	955	AAGAAACTTCGGGTGGAAACTCTGGGCCAGCTGGTAGCAGCTCCAGCCCTGGTGGT	1014		
Qy	1106	ACAAGGCAGTCCCGAGAAGTTACCTATGAGAGCGGGAACTGTGGCCCGGTACATTC	1165		
Db	1015	ACAAGGCAGTCCCGAGAAGTTACCTATGAGAGCGGGAACTGTGGCCCGGTACATTC	1074		
Qy	1166	CATTGGAGAGGGGATTTTACTTACCACGCGGCATTTTGGAGGACGCTACTCACA	1225		
Db	1075	CATTGGAGAGGGGATTTTACTTACCACGCGGCATTTTGGAGGACGCTACTCACA	1134		
Qy	1226	TTCTCAACCTCACCAGGAGTGTCTTAAAGGGATCTCCAGGACAAGAAACATGACATAG	1285		
Db	1135	TTCTCAACCTCACCAGGAGTGTCTTAAAGGGATCTCCAGGACAAGAAACATGACATAG	1194		
Qy	1286	AAGCCAGTGCATGATGAGAGCCACCTCAACAATAGTCTCTTTCANAAACCCACTA	1345		
Db	1195	AAGCCAGTGCATGATGAGAGCCACCTCAACAATAGTCTCTTTCANAAACCCACTA	1254		
Qy	1346	AAATCCTATCTCCAGAGTATTGCTGGGACTATCAGATAGGCTGCCTTCAGATATTAAA	1405		
Db	1255	AAATCCTATCTCCAGAGTATTGCTGGGACTATCAGATAGGCTGCCTTCAGATATTAAA	1314		
Qy	1406	GTGTCAAGGTAGCTTGGCAGACAAAGAGTATAATTTGGTTAGAAATAATCTCTGACTTC	1465		
Db	1315	GTGTCAAGGTAGCTTGGCAGACAAAGAGTATAATTTGGTTAGAAATAATCTCTGACTTC	1374		
Qy	1466	AAATGTGATGGAACCTTGACACTATTCTT	1495		
Db	1375	AAATGTGATGGAACCTTGACACTATTCTT	1404		
RESULT 10					
PIGAL3G					
LOCUS					
DEFINITION Sus scrofa (clone pPGT-3) alpha 1,3-galactosyltransferase mRNA,					
complete cds.					
ACCESSION L36535					
VERSION L36535.1 GI:609566					
KEYWORDS alpha-1,3-galactosyltransferase.					
SOURCE Sus scrofa spleen and liver cDNA to mRNA.					
ORGANISM Sus scrofa					
REFERENCE 1 (bases 1 to 1343)					
AUTHORS Sandrin,M.S., Dabkowski,P.L., Henning,M.M., Mouhtouris,E. and McKenzie,I.F.C.					
TITLE Mammalia; Euthera; Cetartiodactyla; Suina; Suidae; Sus.					
JOURNAL Xenotransplantation (1994) In press					
FEATURES					
source					
1..1343					
/organism="Sus scrofa"					
/db_xref="taxon:9823"					
/clones="pPGT-3"					
/tissue_type="spleen and liver"					
/note="(vector lambda gt10) (vector lambda gt11)"					
1..49					
5'UTR					
mRNA					
CDS					
1..1343					
50..1129					
/codon_start=1					
/product="alpha-1,3-galactosyltransferase"					
/protein_id="AAA58775.1"					
/db_xref="GI:609567"					
/translation="MNVGRVLSMLLVSTVMVWFVEY INRNPEVGSQAORWFFPSH					
FNNGTHSYHEEDAIGNEKQRDNRELPLVDVFNPKRPEVVTITRKAPVWVEG					
TYNRVLDNTYAKQITVGLTFVAVGRYDIEHLPEFLISANTYPMVGHKVFYIMVDD					
ISRMPLJELGPLRSKFVEIKSRKMDI SMRMKTI GEHLAIHQHEVDFLFCIDVD					
QVQNFVETLGOSVAOLQAWNYKAHPDEFTYERRKSAAY IPFGGDFYHAAIFG					
GTPQVNLITQECPKGILODKENDIEAEWHDESHLNKYFLLNKPTK ILSPEYCDYHI					
GMSVDIRIVKIAWKKEYNLRNNI"					



3'UTR									
BASE COUNT		1130	a	281	c	324	g	339	t
ORIGIN									
Query Match 48.4%; Score 726.6; DB 4; Length 1343;									
Best Local Similarity 78.8%; Pred. No. 3e-184;									
Matches 893; Conservative 0; Mismatches 234; Indels 6; Gaps 2;									
Qy	354	CTCAAGATCTGGATACAGCAGAGAAAATAATGAATGTCAAGGGAAAGTAATCTCTGTGTGAT	413						
Db	22	CCCAGCTTCTGCCGATCAGGAGAAAATAATGAATGTCAAGGAAGAGTGTCTGTCAAT	81						
Qy	414	GCTGATTGTCTCAACCGTGGTGTCTGTGTTTGGGAATATGTCAACAGAATTCAGAGGT	473						
Db	82	GCTGCTTGTCTCAACTGTAATGGTGTGTGTTTGGGAATACATCAACAGAACCCAGAAGT	141						
Qy	474	TGTTGAGACAGATGGCAGAGGACTGGTGTCTCCCAAGCTGGTTTAAAAATGGGACCCA	533						
Db	142	TGG---CAGCAGTGTCTCAGAGGGGCTGGTGGTTTCCGAGCTGGTTTAAACAAATGGGACTCA	198						
Qy	534	CAGTTATCAAGAAGACAACTAGAACGCGGAGAGAAAAGG---GTAGAAATGGAGATCG	590						
Db	199	CAGTTACCAGAGAGAACGCTATAGCGAACGAAAGGAACAAAGAAAGAGACAA	258						
Qy	591	CATTGAAGAGCCTCAGCTATGGGACTGGTTCAATCCAAAGAACCGCCGGATGTTTGGAC	650						
Db	259	CAGAGGAGAGCTCCCGCTAGTGGACTGGTTTAACTCTGAGAAACGCCAGAGTCTGTGAC	318						
Qy	651	AGTGACCCCTGGAAAGCGCGGATGTGTGGGAAGGCACTTATGACACAGCTCTGCTGGA	710						
Db	319	CATAACCAAGATGGAAGGCTCCAGTGGTATGGGAAGGCATTTACAACAGAGCCGCTTTAGA	378						
Qy	711	AAAGTACTAGCCACACAGAAACTCACTGTGGGCTCAGACGTGTTGCTGTGGAAAGTA	770						
Db	379	TAATTATTATGCCAAGAGAAATACCGTGGGCTTGACGGTGTTCGCTGTCSGAAGATA	438						
Qy	771	CATTGAGCATTTACTTGAAGACTTCTGGAGTCTGTGACATGTACTTTCATGTTGGCCA	830						
Db	439	CATTGAGCATTTACTTGGAGGAGTCTTAAATCTGCAAAATACATCTCATGTTGGCCA	498						
Qy	831	TCGGGTGATATTTTACGTCATGATAGACGACACTCCCGGATCCCTGTGTCGACCTGAA	890						
Db	499	CAAGTGCATCTTTTACATCATGTTGGATGATATCTCCAGGATGCTTTGATAGAGCTGG	558						
Qy	891	CCCTCTACATTTCTTACAAGTCTTTGAGATCAGTCTGTGAGAAGAGTGGCAGGATATCAG	950						
Db	559	TCCTCTCGTTCCTTTAAAGTGTGTGAGATCAAGTCCGAGAGAGTGGCAAGACATCAG	618						
Qy	951	CATGATCGGCATGAAGACCATTTGGGAGCACATCTCTGGCCACATCCAGCACGAGGTGGA	1010						
Db	619	CATGATCGGCATGAAGACCATCGGGAGCACATCTCTGGCCACATCCAGCACGAGGTGGA	678						
Qy	1011	CTTCTCTCTGTCATGACGTGGATCAAGTCTTTTCAAGACAACTTCCGGGTGGAACTCT	1070						
Db	679	CTTCTCTCTGTCATGACGTGGATCAAGTCTTCCAAACAACTTTGGGTGGAGACCT	738						
Qy	1071	GGGCCAGCTGTGACACAGCTTCCAGGCTGTGTGTGTAAGGCCAGTCCCGAGAGTTTCAC	1130						
Db	739	GGGCCAGCTGTGCTACAGCTACAGGCTGTGTGTGTAAGGCCACATCTCAGCAGTTTCAC	798						
Qy	1131	CTATGAGAGGGGGAACCTGTGCGCGCGGTACATTTCCATTCGGAGAGGGGATTTTTACTA	1190						
Db	799	CTAGGAGAGGGGGAAGGATCCGAGGCTACATTTCCGCTTTGGCCAGGGGATTTTTATTA	858						
Qy	1191	CCAGCGGGCCATTTTTCAGAGAACGCTTACTACATTTCTCAACCTCACAGGAGTGCCT	1250						
Db	859	CCAGCGAGCCATTTTGGGGGAACCCACTCAGGTGTTCTAAACATCTCAGGAGTGCCT	918						
Qy	1251	TAAGGGGATCCTCCAGGACAGAAACATGACATAGAGCCAGTGGCATGATGAGAGCCA	1310						
Db	919	CAAGGGNAATCCTCCAGGACAGGNAATGACATAGAGCCAGTGGCATGATGAAAGCCA	978						

Qy	1311	CCTCAACAAATACCTCTCTTTTCAACAAACCCACTAAAAATCCTATCTCCAGAGTATTCGCTG	1370																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																										</
----	------	---	------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----





QY 1306 AGCCACCTCAACAAATACCTTCCCTTTTCAACAAACCACATAAAATCCTATCTCCAGAGTAT 1365  
|||||  
Db 935 AGCCACCTCAACAAATACCTTCCCTTTTCAACAAACCACATAAAATCCTATCTCCAGAGTAT 994  
QY 1366 TGCCTGGACTATCAGATAGGCGCTTCCAGTATTTAAAGTGTCAAGGTAGCTTGGCAG 1425  
|||||  
Db 995 TGCCTGGACTATCAGATAGGCGCTTCCAGTATTTAAAGTGTCAAGGTAGCTTGGCAG 1054  
QY 1426 ACAAAGAGTATANTTTGGTTAGAAATAATGCTGACTTCAAAATCTGATGGAACCTTGA 1485  
|||||  
Db 1055 ACAAAGAGTATANTTTGGTTAGAAATAATGCTGACTTCAAAATCTGATGGAACCTTGA 1114  
QY 1486 CACTATTCTT 1495  
|||||  
Db 1115 CACTATTACT 1124  
RESULT 13  
AXI28258  
LOCUS AXI28258 3240 bp DNA linear PAT 15-MAY-2001  
DEFINITION Sequence 25 from Patent WO0130992.  
ACCESSION AXI28258  
VERSION AXI28258.1 GI:14134779  
KEYWORDS  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 3240)  
AUTHORS Koike,C.  
TITLE \_g(a)1-3 galactosyltransferase gene and promoter  
JOURNAL Patent: WO 0130992-A 25 03-MAY-2001;  
UNIV. PITTSBURGH OF THE COMMONWEALTH SYSTEM OF HIGHER EDUCATION  
(US)  
FEATURES  
source Location/Qualifiers  
1..3240  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
intron 1..369  
misc\_feature 370..3010  
/note="exon 9"  
terminator 1088..1090  
3'UTR 1091..3010  
intron 3011..3240  
BASE COUNT 901 a 694 c 727 g 878 t 40 others  
ORIGIN  
Query Match 48.3%; Score 725.2; DB 6; Length 3240;  
Best Local Similarity 99.6%; Pred. No. 8.1e-184;  
Matches 727; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 766 AAGTACATTGAGGATTACTTAGAAGACTTTCTGGAGCTTCGTGACATGTACTTCATGGTT 825  
Db 395 AGGTACATTGAGGATTACTTAGAAGACTTTCTGGAGCTTCGTGACATGTACTTCATGGTT 454  
QY 826 GGCCATCGGGTCAATATTTACGTCATGATAGACGACACCTCCGGGATGCTGTCGTGCAC 885  
Db 455 GGCCATCGGGTCAATATTTACGTCATGATAGATGACACCTCCGGGATGCTGTCGTGCAC 514  
QY 886 CTGAACCCCTCACTATCTTACAAAGCTTTTGAGATCAGGCTCTGAGAAGAGGTGGCAGGAT 945  
Db 515 CTGAACCCCTCACTATCTTACAAAGCTTTTGAGATCAGGCTCTGAGAAGAGGTGGCAGGAT 574  
QY 946 ATCAGCATGATGGCATGAAGACCATTTGGGGAGCAGCATCTCGGCCACATCCAGCACGAG 1005  
Db 575 ATCAGCATGATGGCATGAAGACCATTTGGGGAGCAGCATCTCGGCCACATCCAGCACGAG 634  
QY 1006 GTCGACTTCTCTTCGATGACGCTGGATCAAGTCTTTCAAGACAACTTCGGGGTGGAA 1065  
Db 635 GTCGACTTCTCTTCGATGACGCTGGATCAAGTCTTTCAAGACAACTTCGGGGTGGAA 694

QY 1066 ACTCTGGGGCCAGCTGCTAGCACAGCTCCAGGCTGCTGGTACAAAGCCAGTCCCGAGAAG 1125  
|||||  
Db 695 ACTCTGGGGCCAGCTGCTAGCACAGCTCCAGGCTGCTGGTACAAAGCCAGTCCCGAGAAG 754  
QY 1126 TTCACCTATTCAGAGGGCGGAACTGTCGGCGCGGTACATTTCCCATTTCCGAGAGGGGGATTTT 1185  
|||||  
Db 755 TTCACCTATTCAGAGGGCGGAACTGTCGGCGCGGTACATTTCCCATTTCCGAGAGGGGGATTTT 814  
QY 1186 TACTACACGCGGCCATTTTGGAGGAAGCCCTACTTCACATTTCTCAACCTTCACACGGGAG 1245  
|||||  
Db 815 TACTACACGCGGCCATTTTGGAGGAAGCCCTACTTCACATTTCTCAACCTTCACACGGGAG 874  
QY 1246 TGCCTTAAAGGGATCTCCAGGACAAAGAACATGACATAGAACCCAGTGGCATGATGAG 1305  
|||||  
Db 875 TGCCTTAAAGGGATCTCCAGGACAAAGAACATGACATAGAACCCAGTGGCATGATGAG 934  
QY 1306 AGCCACCTCAACAAATACCTTCCCTTTTCAACAAACCACATAAAATCCTATCTCAGAGTAT 1365  
|||||  
Db 935 AGCCACCTCAACAAATACCTTCCCTTTTCAACAAACCACATAAAATCCTATCTCAGAGTAT 994  
QY 1366 TGCCTGGACTATCAGATAGGCGCTTCCAGTATTTAAAGTGTCAAGGTAGCTTGGCAG 1425  
|||||  
Db 995 TGCCTGGACTATCAGATAGGCGCTTCCAGTATTTAAAGTGTCAAGGTAGCTTGGCAG 1054  
QY 1426 ACAAAGAGTATANTTTGGTTAGAAATAATGCTGACTTCAAAATCTGATGGAACCTTGA 1485  
|||||  
Db 1055 ACAAAGAGTATANTTTGGTTAGAAATAATGCTGACTTCAAAATCTGATGGAACCTTGA 1114  
QY 1486 CACTATTCTT 1495  
|||||  
Db 1115 CACTATTACT 1124  
RESULT 14  
A69346  
LOCUS A69346 1092 bp DNA linear PAT 06-MAY-1999  
DEFINITION Sequence 3 from Patent WO9802453.  
ACCESSION A69346  
VERSION A69346.1 GI:4760173  
KEYWORDS  
SOURCE unidentified.  
ORGANISM unidentified.  
REFERENCE 1 (bases 1 to 1092)  
AUTHORS Gourlet,P., Robberecht,P., Vandermeers,A. and Woelbroeck,M.  
TITLE PEPTIDIC LIGANDS HAVING A HIGHER SELECTIVITY FOR THE VIP1 RECEPTOR  
JOURNAL THAN FOR THE VIP2 RECEPTOR  
Patent: WO 9802453-A 3 22-JAN-1998;  
UNIV BRUXELLES (BE)  
FEATURES  
source Location/Qualifiers  
1..1092  
/organism="unidentified"  
/db\_xref="taxon:32644"  
CDS <1..1092  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAB42481.1"  
/db\_xref="GI:4760174"  
/translation="NSKIMNVKGRVLSMLLYSTVNVVFWVEYINRNPEVGSQAQRGW  
PFSWFNNGTHSHEEDAIQNEKEQRKEDNKGELPLVDKFNPEKRPVVVITRWKAPV  
VWEGTYNRAVLNDNYAKQITVGLTVLAVGRYIEHYLEEFLLSANTYFMVGHVPIFYI  
MVDDISRMPLIELGRLSPFKVPEIKSEKRWDISMMRMKTIGEHILAHIQHVDPLFC  
MDVDQVFNNGFVETLGOSVAOLOAWYKAHPDEFTYERRKESAAYIPFGQGDFFYYHA  
AIFGCTPTQVLNITQCEPKGILQDKENDLEAEWHDESHLNKFLNKPTKILSPYCVH  
DTHICMSVDIRIVKIAWQKRNLYNRNNI"  
BASE COUNT 325 a 228 c 279 g 260 t  
ORIGIN  
Query Match 47.1%; Score 706; DB 6; Length 1092;  
Best Local Similarity 79.3%; Pred. No. 9.9e-179;  
Matches 864; Conservative 0; Mismatches 220; Indels 6; Gaps 2;

QY 375 GAAATAATGAATGTCAGGGAAAGTAATCCCTGTTGATGCTGATGCTCTCAACCGTGGT 434  
Db 6 GAAATAATGAATGTCAGGAAGAGTGGTCTGTCAATGCTGCTCTCAACTGTAAT 65  
QY 435 TGTGCTGCTTTTGGGAATATGTCACACAAATCCAGAGAGTTGGTCAGACAGATGGCAGAA 494  
Db 66 GGTGTGCTTTTGGGAATATGTCACACAAATCCAGAGAGTTGGTCAGAGTGGTCAGAG 122  
QY 495 GGACTGTGCTTTCCCAAGCTGGTTTAAATAATGGGACCCACAGTTATCAAGAACAACACGT 554  
Db 123 GGCTGTGCTTTCCGAGCTGGTTTACAAATGGGACTCACAGTTACCAACAGAGAGAGA 182  
QY 555 AGAGGACCGAGAGAAAGG---GTAGAAATGGAGATCGCATTTGAAGAGCCTCAGCTATG 611  
Db 183 CGCTATAGGCAAGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 242  
QY 612 GGACTGTGTTCAATCCAAAGACCGCCGGATGTTTACACAGTACACAGTACACAGTAC 671  
Db 243 GGCTGTGCTTTCCGAGCTGGTTTACAAATGGGACTCACAGTTACCAACAGAGAGAGA 302  
QY 672 GATTGTGTGGGAAGGCACTTATGACACAGCTCTGCTGGAAAGTACTACGCCACACAGAA 731  
Db 303 AGTGGTATGGAAGGCACTTACAAACAGAGCGCTTATAGATAATATTATGCCAAACAGAA 362  
QY 732 ACTCACTGTGGGGCTGACAGTGTGTTGCTGTGGGAAGTACATTTGAGCATTTACTTAGA 791  
Db 363 AATTACCGTGGGCTGACGCTGTTGCTGTGGGAAGTACATTTGAGCATTTACTTTGAGGA 422  
QY 792 CTCTTCGAGCTGCTGACATGACTTTCATGTTGGCCATCGGGTCATATTTTACGTCAT 851  
Db 423 GTTCTTAAATCTGCAAAATACATCTTTCATGTTGGCCCAAAAGTCACTTTTACATCAT 482  
QY 852 GATAGAGGACACCTCCCGGATGCTGCTGACCTGCACTGAAACCCCTACATTTCTTACAA 911  
Db 483 GGTGGATGATATCTCCAGGATGCTTTGATAGAGCTGGGCTCTGCGCTCTTAAAGT 542  
QY 912 CTCTTGTAGATCAGTCTGAGAGAGTGGCAGATATCAGCATGATGCGCATGAAGACCAT 971  
Db 543 GTTTGAGATCAAGTCCGAGAGAGTGGCAAGCATCAGCATGATGCGCATGAAGACCAT 602  
QY 972 TGGGAGCACATCTGGCCACATCCAGCAGAGTTCGCTCTCTCTGTCATGGAAGT 1031  
Db 603 CGGGAGCACATCTGGCCACATCCAGCAGAGTTCGCTCTCTCTGTCATGGAAGT 662  
QY 1032 GGATCAAGTCTTTCAAGACAACTTCGGGGTGGAACTCTGGGCGAGCTGTAGCAGCAT 1091  
Db 663 GGATCAGGCTTCCAAACAACTTTGGGGTGGAGACCTTGGCCAGCTGGGCTCAGCT 722  
QY 1092 CCAGGCTGTGTGTCAGAGCCAGTCCGAGAGTTTACCTATGAGAGCGGGNACTGTC 1151  
Db 723 ACAGGCTGTGTGTCAGAGCCATCTGACGAGTTTACCTATGAGAGCGGGNAAGGAGTC 782  
QY 1152 GGCGCGTACATTTCCATTCGAGAGGGGATTTTACTACAGCGCGGCAATTTTGGAGG 1211  
Db 783 CGCAGCTACATTTCCGTTTGGCCAGGGGATTTTATACAGCAGCCATTTTGGGGG 842  
QY 1212 AACCCCTACTACATTTCTCAACCTCACAGGAGTGGTTTAAAGGGATCTCTCCAGGACAA 1271  
Db 843 AACCCCTACTCAGTTTCTAAACATCACTCAGGAGTGGTTTCAAGGGAATCTCTCCAGGACAA 902  
QY 1272 GAACATGACATAGAGCCAGCTGGCATGATGAGAGCCACTCAACAAATACTTCTCTTT 1331  
Db 903 GGAATATGATATAGAGCCAGTGGCATGATGAAGCCATCTAAACAGTATTTTCTCTCT 962  
QY 1332 CAACAAACCCACTAAATCTCTATCCAGAGTATGCTGGGACTATCAGATAGCCCTGCC 1391  
Db 963 CAACAAACCCACTAAATCTTATCCCAAGAAATATCTGCTGGGATATCATATAGGCAATGTC 1022  
QY 1392 TTCAGATATTTAAAGTGTCAAGGTAGTGGCAGACAAAAGATATATTTTGGTTAGAAA 1451  
Db 1023 TGTGGATATTTAGGATTTGTCAGATAGCTTGGCAGAAAAGAGATATATTTTGGTTAGAAA 1082  
QY 1452 TAATGTCCTGA 1461

Db 1083 TAACATCTGA 1092

RESULT 15

A72023

LOCUS

DEFINITION

Sequence 3 from Patent WO9803653.

ACCESSION

A72023.1

VERSION

A72023.1

KEYWORDS

unidentified.

unidentified

unclassified.

REFERENCE

1 (bases 1 to 1092)

AUTHORS

Pourcel,C., Soullilou,J. and Vanhove,B.

TITLE

METHOD FOR PREPARING TRANSGENIC NON-HUMAN MAMMALIAN ORGANS FOR

TRANSPLANTATION TO HUMANS, AND NUCLEOTIDE SEQUENCES THEREFOR

JOURNAL

Patent: WO 9803653-A 29-JAN-1998;

INST NAT SANTE RECH MED (FR)

Other publication FR 2751346 19980123.

COMMENTS

Location/Qualifiers

1..1092

/organism="unidentified"

/db\_xref="taxon:32844"

<1..1092

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAB42704.1"

/db\_xref="GI:4808050"

BASE COUNT

325 a 228 c 279 g 260 t

Query Match

Best Local Similarity

Matches 864; Conservative

0; Mismatches 229;

Indels 6; Gaps 2;

Score 706; DB 6; Length 1092;

Pred. No. 9.9e-179;

QY 375 GAAATAATGAATGTCAGGGAAAGTAATCCCTGTTGATGCTGATGCTCTCAACCGTGGT 434

Db 6 GAAATAATGAATGTCAGGAAGAGTGGTCTGTCAATGCTGCTCTCAACTGTAAT 65

QY 435 TGTGCTGCTTTTGGGAATATGTCACACAAATTCACAGAGTTGGTGAGAACAGATGGCAGAA 494

Db 66 GGTGTGCTTTTGGGAATATGTCACACAAATTCACAGAGTTGGTGAGAACAGATGGCAGAA 122

QY 495 GGACTGTGCTTTCCCAAGCTGGTTTAAATAATGGGACCCACAGTTATCAAGAACAACACGT 554

Db 123 GGCTGTGCTTTCCGAGCTGGTTTACAAATGGGACTCACAGTTACCAACAGAGAGAGA 182

QY 555 AGAGGACCGAGAGAAAGG---GTAGAAATGGAGATCGCATTTGAAGAGCCTCAGCTATG 611

Db 183 CGCTATAGGCAAGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 242

QY 612 GGACTGTGTTCAATCCAAAGACCGCCGGATGTTTACACAGTACACAGTACACAGTAC 671

Db 243 GGCTGTGCTTTCCGAGCTGGTTTACAAATGGGACTCACAGTTACCAACAGAGAGAGA 302

QY 672 GATTGTGTGGGAAGGCACTTATGACACAGCTCTGCTGGAAAGTACTACGCCACACAGAA 731

Db 303 AGTGGTATGGAAGGCACTTACAAACAGAGCGCTTATAGATAATATTATTTAGCAAGAA 362

QY 732 ACTCACTGTGGGCTGACAGTGTGTTGCTGTGGAAAGTACATTCAGCATTTACTTAGAAGA 791

Db 363 AATTACCGTGGGCTGACGCTGTTGCTGTGGGAAGATACATTTGAGCATTTACTTTGAGGA 422

Qy	792	CTTTCTGGAGTCTGCTTGACATGTA	CTTCTATGGTTGGCCATCGGGTCATATTTTACGTCAT	851
Db	423	GTTCCTTAATATCTGCAAAATACAT	TTCTATGGTTGGCCAAAGATCATCTTTTACATCAT	482
Qy	852	GATAGACGACACCTCCCGGATGCCT	GTGTCACCTGAACCTCTACATTCCTTTACAAGT	911
Db	483	GGTGGATGATATCTCCAGGATGCC	TTTGATAGAGCTGGGTCTCTGCGCTCTTTAAGT	542
Qy	912	CTTTTGAGATCAGTCTTGAGAAGAGT	GGCAGGATATCAGCATGATGGCGCATGAAGACCAT	971
Db	543	GTTTTGAGATCAAGTCCGCGAAGAG	GTGGCAGACATCAGCATGATGGCGCATGAAGACCAT	602
Qy	972	TGGGAGACACATCCTGGCCACAT	CACGACAGGTGCACCTTCCTCTTCGTCATGGACGT	1031
Db	603	CGGGAGCACATCCTGGCCACAT	CACGACAGGTGCACCTTCCTCTTCGTCATGGACGT	662
Qy	1032	GGATCAAGTCTTTCAAGACAACT	CTCGGGTGGAACTCTGGGCCAGCTGGGTAGCACAGCT	1091
Db	663	GGATCAGGTTCTCCAAACAACT	TTTGGGTGGAGACCTCGGGCCAGTCGGTGGCTCAGCT	722
Qy	1092	CCAGGCTGTGTGTACAGGCCAGT	CCCGAGAGTTACCTATCAGAGGGCGGGAACGTGTC	1151
Db	723	ACAGGCTGTGTGTACAGGCCAC	TCTGACGAGTTACCTACGAGAGGGGAAGGAGTC	782
Qy	1152	GGCCGGGTACATTTCCATTCGGAG	GGGGATTTTTTACTACCACGCGGCCATTTTTTGGAGG	1211
Db	783	CGCAGCCTACATTTCCCTTTGGC	CAGGGGATTTTTTATTACCACGACCCATTTTTTGGGG	842
Qy	1212	AACGCCCTACTACATTTCTCAAC	CTCACCAGGAGTGCTTTAAGGGGATCTCCAGGACAA	1271
Db	843	AACACCCACTACGTTTCTAAAT	CACCTCAGGAGTGCTTCAAGGGAACTCTCCAGGACAA	902
Qy	1272	GAAACATGACATAGAAGCCCACT	TGGCATGATCAGAGCCACCTCAACAAATACTTCCCTTTT	1331
Db	903	GGAAATGACATAGAAGCCCACT	TGGCATGATGAAGCCCATCTAAACAGTATTTTCCCTTCT	962
Qy	1332	CACAAACCCCACTAAATCCTAT	CTCCAGAGTATTTGCTGGGCACTATCAGATAGGCCCTGCC	1391
Db	963	CAACAAACCCCACTAAATCTTAT	TCCCGAGAATACTGCTGGGATTTATCATATAGGCATGTC	1022
Qy	1392	TTCAGATATTTAAAGTGTCAAG	GTAGCTTGGCGACACAAAGAGTATATTTTGGTTAGAAA	1451
Db	1023	TGTGGATATTTAGGATTTGTC	CAGATACCTTGGCGAAGAAAGAGTATATTTTGGTTAGAAA	1082
Qy	1452	TAATGTCGTA	1461	
Db	1083	TAACATCTGA	1092	

Search completed: May 17, 2002, 19:27:30  
Job time: 12075 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 16:14:02 ; Search time 70.18 Seconds  
(without alignments)  
5250.075 Million cell updates/sec

Title: US-09-863-475A-3

Perfect score: 1500

Sequence: 1 CTTGCCCTTGAGACTTTC.....CTTGACACTATTCTAACCA 1500

Scoring Table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued\_Patents\_NA:\*
- 1: /cgn2\_6/ptodata/2/lna/5A\_COMB.seq:\*
  - 2: /cgn2\_6/ptodata/2/lna/5B\_COMB.seq:\*
  - 3: /cgn2\_6/ptodata/2/lna/6A\_COMB.seq:\*
  - 4: /cgn2\_6/ptodata/2/lna/6B\_COMB.seq:\*
  - 5: /cgn2\_6/ptodata/2/lna/PCTUS\_COMB.seq:\*
  - 6: /cgn2\_6/ptodata/2/lna/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1500	100.0	1500	1	US-07-914-281-3
2	1500	100.0	1500	1	US-08-393-246-3
3	1500	100.0	1500	1	US-08-525-058A-3
4	1500	100.0	1500	2	US-08-696-731-3
5	1500	100.0	1500	4	US-09-042-531-3
6	1454.6	97.0	1500	5	PCT-US91-00899-2
7	1408.4	93.9	3450	2	US-08-378-617A-9
8	1005.8	67.1	1131	2	US-08-704-548-3
9	726.6	48.4	1423	1	US-08-214-580A-2
10	726.6	48.4	1423	5	PCT-US95-07534-5
11	702.8	46.9	1140	2	US-08-704-548-1
12	684.2	45.6	1269	3	US-08-621-700-1
13	684.2	45.6	1269	5	PCT-US95-03940-1
14	684.2	45.6	1412	2	US-08-378-617A-7
15	642.4	42.8	1828	2	US-08-378-617A-8
16	583.2	38.9	1353	1	US-08-214-580A-1
17	241.6	16.1	826	4	US-08-853-774-1
18	240	16.0	1062	1	US-07-752-101A-34
19	233.6	15.6	826	4	US-08-853-774-2
20	232	15.5	826	4	US-08-853-774-3
21	232	15.5	826	4	US-08-853-774-4
22	188.6	12.6	585	1	US-07-752-101A-66
23	83.4	5.6	463	1	US-07-752-101A-67
24	54.6	3.6	7218	1	US-08-232-463-14
25	36.2	2.4	2469	1	US-07-997-133-2
26	36.2	2.4	2469	1	US-08-459-296-1
27	36.2	2.4	2469	5	US-07-997-133-2

28	36.2	2.4	2662	2	US-08-451-822A-14	Sequence 14, Appl
29	36.2	2.4	2662	4	US-08-323-430-14	Sequence 14, Appl
30	36.2	2.4	2733	1	US-08-371-001-14	Sequence 14, Appl
31	36.2	2.4	2733	5	PCT-US96-00331-14	Sequence 14, Appl
32	35.8	2.4	282	2	US-08-716-942-18	Sequence 18, Appl
33	35.2	2.3	289	4	US-09-007-005-17	Sequence 17, Appl
34	35.2	2.3	289	4	US-09-244-796-17	Sequence 17, Appl
35	35.2	2.3	547	3	US-09-188-930-14	Sequence 14, Appl
36	34.8	2.3	1416	3	US-08-911-853-3	Sequence 3, Appl1
37	34.8	2.3	1416	4	US-09-479-409-3	Sequence 3, Appl1
38	34.8	2.3	1416	4	US-09-479-453-3	Sequence 3, Appl1
39	34.8	2.3	4377	3	US-08-911-853-28	Sequence 28, Appl
40	34.8	2.3	4377	4	US-09-479-409-28	Sequence 28, Appl
41	34.8	2.3	4377	4	US-09-479-453-28	Sequence 28, Appl
42	34.2	2.3	360	5	PCT-US93-05703-1	Sequence 1, Appl1
43	33.6	2.2	1001	3	US-08-705-771-6	Sequence 6, Appl1
44	33.4	2.2	1179	2	US-08-465-794-4	Sequence 4, Appl1
45	33.4	2.2	1179	3	US-09-049-813-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1  
US-07-914-281-3  
; Sequence 3, Application US/07914281  
; Patent No. 5324663  
; GENERAL INFORMATION:  
; APPLICANT: LOWE, JOHN B.  
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
; OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATIER & NEUSTADT,  
; P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/914,281  
; FILING DATE: 19920720  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lavalleye, Jean-Paul M. P.  
; REGISTRATION NUMBER: 31,451  
; TELEPHONE/DOCKET NUMBER: 2363-060-55  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)521-4500  
; TELEFAX: (703)486-2347  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1500 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
; ANTI-SENSE: NO  
; US-07-914-281-3

Query Match 100.0%; Score 1500; DB 1; Length 1500;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 CTTTCCCTTGTAGACTCTTCTTGAATGAGAAGTACCGATTCTCTGCTGAAGACCTCGCGCT 60
Db 1 CTTTCCCTTGTAGACTCTTCTTGAATGAGAAGTACCGATTCTCTGCTGAAGACCTCGCGCT 60
Qy 61 CTCAGGCTCTGGGAGTTGGAAACCTGTACCTTCCTTCCCTCTGCTGAGCCGCTCGCTCTT 120
Db 61 CTCAGGCTCTGGGAGTTGGAAACCTGTACCTTCCTTCCCTCTGCTGAGCCGCTCGCTCTT 120
Qy 121 AGGAGGCCAGAGCTCGACAGAACTCGTTCCTTGTGCTTGTGCTTGGAGGGAACACAG 180
Db 121 AGGAGGCCAGAGCTCGACAGAACTCGTTCCTTGTGCTTGTGCTTGGAGGGAACACAG 180
Qy 181 CTGACGATGAGGCTGACTTTGAATCAAGAGATCTGTTACCCAGTCTCTCGTGAATTA 240
Db 181 CTGACGATGAGGCTGACTTTGAATCAAGAGATCTGTTACCCAGTCTCTCGTGAATTA 240
Qy 241 AGGCTGTACTACATTTGCTGGAGACCTAAGATTTTCATGATCACTATGCTTCAAGATCTC 300
Db 241 AGGCTGTACTACATTTGCTGGAGACCTAAGATTTTCATGATCACTATGCTTCAAGATCTC 300
Qy 301 CATCTCAACAGATCTCCATGTCGAAGTCCAAAGTCAGAAACAAGTCTTCCATCTCAAGA 360
Db 301 CATCTCAACAGATCTCCATGTCGAAGTCCAAAGTCAGAAACAAGTCTTCCATCTCAAGA 360
Qy 361 TCTGGATCACAGGAGAAATTAATGAATGTCAGGGAAGAAAGTAAATCCTGTTGATGCTGATT 420
Db 361 TCTGGATCACAGGAGAAATTAATGAATGTCAGGGAAGAAAGTAAATCCTGTTGATGCTGATT 420
Qy 421 GTCTCAACCGTGTGCTGCTTTTGGGAATATGTCACAGAAATTCAGAGGTTGGTGAG 480
Db 421 GTCTCAACCGTGTGCTGCTTTTGGGAATATGTCACAGAAATTCAGAGGTTGGTGAG 480
Qy 481 AACAGATGGCAGAGGACTGCTGCTCCAAAGTGGTTTAAAAATGGGACCCACAGATTAT 540
Db 481 AACAGATGGCAGAGGACTGCTGCTCCAAAGTGGTTTAAAAATGGGACCCACAGATTAT 540
Qy 541 CAAGAAGACAACTGTAAGGACGAGAGAAAGGTAAGAAATGGAGATCGCATTTGAAGAG 600
Db 541 CAAGAAGACAACTGTAAGGACGAGAGAAAGGTAAGAAATGGAGATCGCATTTGAAGAG 600
Qy 601 CCTCAGTATGGGACTGTTCAATCCAAAGAACCCCGGATGTTTGCACAGTCAACCCG 660
Db 601 CCTCAGTATGGGACTGTTCAATCCAAAGAACCCCGGATGTTTGCACAGTCAACCCG 660
Qy 661 TGAAGGCCGCGATTGCTGGGAAGGCACTTATGACACAGCTCTGCTGGAAAGTACTAC 720
Db 661 TGAAGGCCGCGATTGCTGGGAAGGCACTTATGACACAGCTCTGCTGGAAAGTACTAC 720
Qy 721 GCCACAGAAACTCAGTGTGGGCTGACAGTGTGCTGTGGGAAGTACATTTGAGCAT 780
Db 721 GCCACAGAAACTCAGTGTGGGCTGACAGTGTGCTGTGGGAAGTACATTTGAGCAT 780
Qy 781 TACTTAGAAGACTTCTGGAGTCTGCTGACATGTACTTCAATGTTGGCCATCGGGTCATA 840
Db 781 TACTTAGAAGACTTCTGGAGTCTGCTGACATGTACTTCAATGTTGGCCATCGGGTCATA 840
Qy 841 TTTTACGTCATGATAGACGACACCTCCCGGATGCTGCTGTCGACCTGAACCCCTCTACAT 900
Db 841 TTTTACGTCATGATAGACGACACCTCCCGGATGCTGCTGTCGACCTGAACCCCTCTACAT 900
Qy 901 TCCCTTAAAGTCTTTGAGATCAGTCTGAGAAAGGTTGGCAGGATATCAGCATGATCGGC 960
Db 901 TCCCTTAAAGTCTTTGAGATCAGTCTGAGAAAGGTTGGCAGGATATCAGCATGATCGGC 960
Qy 961 ATGAAGACCATTTGGGAGCAGATCCTGCCCCACATCCAGCAGAGGTCGACTTCTCTTC 1020
Db 961 ATGAAGACCATTTGGGAGCAGATCCTGCCCCACATCCAGCAGAGGTCGACTTCTCTTC 1020
Qy 1021 TGCATGGACGTGGATCAAGTCTTTCAAGACAACTTCGGGGTGGAAACTCTCGGCCAGCTG 1080
Db 1021 TGCATGGACGTGGATCAAGTCTTTCAAGACAACTTCGGGGTGGAAACTCTCGGCCAGCTG 1080
```

## RESULT 2

```
US-08-393-246-3
; Sequence 3, Application us/08393246
; Patent No. 5595900
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; OF OLIGOSACCHARIDES, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESS: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,246
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,433
; FILING DATE: 30-MAR-1994
; APPLICATION NUMBER: US 07/914,281
; FILING DATE: 20-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1500 base pairs
```

TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: cDNA  
ANTI-SENSE: NO  
US-09-393-246-3

Query Match 100.0%; Score 1500; DB 1; Length 1500;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CCCTCCCTTGTAGACTCTTCTGGAATGAGAAGTACCGGATTCCTGGAAGACCTCGCGCT	60
DB	1	CCCTCCCTTGTAGACTCTTCTGGAATGAGAAGTACCGGATTCCTGGAAGACCTCGCGCT	60
QY	61	CTCAGGCTCTGGAGTTGGAAACCTGTACCTTCTTCCCTGCTGAGCCCTGCCTCCTT	120
DB	61	CTCAGGCTCTGGAGTTGGAAACCTGTACCTTCTTCCCTGCTGAGCCCTGCCTCCTT	120
QY	121	AGCAGGCCAGAGCTCGACAGAACTCGGTTTGTCTTGTCTTGGAGGGAACACAG	180
DB	121	AGCAGGCCAGAGCTCGACAGAACTCGGTTTGTCTTGTCTTGGAGGGAACACAG	180
QY	181	CTGACGATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCAGTCTCCTGGAATTA	240
DB	181	CTGACGATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCAGTCTCCTGGAATTA	240
QY	241	AGCCCTGACTACATTTGCCCTGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCTC	300
DB	241	AGCCCTGACTACATTTGCCCTGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCTC	300
QY	301	CATGTCAACAAGATCTCCATGTCAAGATCCAAGTCAAGAAACAAGTCTTCCATCCTCAAGA	360
DB	301	CATGTCAACAAGATCTCCATGTCAAGATCCAAGTCAAGAAACAAGTCTTCCATCCTCAAGA	360
QY	361	TCTGGATCACAGAGAAATAATGAATGTCAAGGAAAGTAATCCTGTGTGATGCTGATT	420
DB	361	TCTGGATCACAGAGAAATAATGAATGTCAAGGAAAGTAATCCTGTGTGATGCTGATT	420
QY	421	GTCTCAACCTGTGTGCTGTTTGGGAATATGTCAACAGAAATTCAGAGGTTGGTGAG	480
DB	421	GTCTCAACCTGTGTGCTGTTTGGGAATATGTCAACAGAAATTCAGAGGTTGGTGAG	480
QY	481	AACAGATGGCAGAGGACTGGTGTCCCAAGCTGGTTTAAAAATGGGACCCACAGTTAT	540
DB	481	AACAGATGGCAGAGGACTGGTGTCCCAAGCTGGTTTAAAAATGGGACCCACAGTTAT	540
QY	541	CAAGAACAACGTAGAGGACGAGAGAAAGGTAGAAATGGAGATCGCATTTGAAGAG	600
DB	541	CAAGAACAACGTAGAGGACGAGAGAAAGGTAGAAATGGAGATCGCATTTGAAGAG	600
QY	601	CCTCAGCTATGGGACTGGTTCAATCCAAAGAACCGCCCGGATGTTTGGACAGTGACCCCG	660
DB	601	CCTCAGCTATGGGACTGGTTCAATCCAAAGAACCGCCCGGATGTTTGGACAGTGACCCCG	660
QY	661	TGGAAGCGCGGATGTGTGGAGGACACTTATGACACAGCTCTGCTGGAAAAGTACTAC	720
DB	661	TGGAAGCGCGGATGTGTGGAGGACACTTATGACACAGCTCTGCTGGAAAAGTACTAC	720
QY	721	GCCACAGAAACTACTGTGGGCTGACAGTGTCTGCTGGGAAAGTACATTTGACAT	780
DB	721	GCCACAGAAACTACTGTGGGCTGACAGTGTCTGCTGGGAAAGTACATTTGACAT	780
QY	781	TACTTAAAGACTTCTTGGAGTCTGCTGACATGTACTTTCATGTTGGCCATCGGGTATA	840
DB	781	TACTTAAAGACTTCTTGGAGTCTGCTGACATGTACTTTCATGTTGGCCATCGGGTATA	840
QY	841	TTTTACGTCATGATAGACGACCTCCCGGATGCTGCTGTCACCTGAAACCTCTACAT	900
DB	841	TTTTACGTCATGATAGACGACCTCCCGGATGCTGCTGTCACCTGAAACCTCTACAT	900
QY	901	TCCTTACAAGTCTTTGAGATCAGGCTGTGAGAAGAGGTGGCAGGATATCAGCATGATCGCG	960

DB	901	TCCTTACAAGTCTTTGAGATCAGGTCCTGAGAAGAGGTGGCAGGATATCAGCATGATCGCG	960
QY	961	ATGAAGACCATTTGGGAGCAGCATCTCTGGCCACATCCAGCACGAGTCTGCTCTCTTC	1020
DB	961	ATGAAGACCATTTGGGAGCAGCATCTCTGGCCACATCCAGCACGAGTCTGCTCTCTTC	1020
QY	1021	TGCATGGACCTGGATCAAGTCTTTTCAAGACAACCTTCGGGTGGAAACTCTGGGCCAGCTG	1080
DB	1021	TGCATGGACCTGGATCAAGTCTTTTCAAGACAACCTTCGGGTGGAAACTCTGGGCCAGCTG	1080
QY	1081	GTAGCAGAGCTCCAGGCTGGTGTACAGAGCCAGTCCCGAGAGTTTCACTTATGAGAGG	1140
DB	1081	GTAGCAGAGCTCCAGGCTGGTGTACAGAGCCAGTCCCGAGAGTTTCACTTATGAGAGG	1140
QY	1141	CGGAACTGTGCGCGCTGATTTCCATTCGGAGAGGGGATTTTACTACCACGCGGCC	1200
DB	1141	CGGAACTGTGCGCGCTGATTTCCATTCGGAGAGGGGATTTTACTACCACGCGGCC	1200
QY	1201	ATTTTGGAGAACGCTTACTCATTTCTCAACCTCACAGGGAGTCTTTAAGGGGATC	1260
DB	1201	ATTTTGGAGAACGCTTACTCATTTCTCAACCTCACAGGGAGTCTTTAAGGGGATC	1260
QY	1261	CTCAGGACAAGAACATGACATAGAACCCAGTGGCATGATGAGGCCACCTCAACAAA	1320
DB	1261	CTCAGGACAAGAACATGACATAGAACCCAGTGGCATGATGAGGCCACCTCAACAAA	1320
QY	1321	TACTTCTCTTTTCAACAAACCCACTTAAATCCTATCTCCAGAGTATTTCTGGGACTATCAG	1380
DB	1321	TACTTCTCTTTTCAACAAACCCACTTAAATCCTATCTCCAGAGTATTTCTGGGACTATCAG	1380
QY	1381	ATAGGCTGCTCTCAGATATTTAAAGTGTCAAGGTAGCTTGGCAGACAAAGAGTATAT	1440
DB	1381	ATAGGCTGCTCTCAGATATTTAAAGTGTCAAGGTAGCTTGGCAGACAAAGAGTATAT	1440
QY	1441	TTGGTTAGAAATATGTCTGACTTCAAAATTTGTATGTAAGAACTTGACACTATTTCTAACCA	1500
DB	1441	TTGGTTAGAAATATGTCTGACTTCAAAATTTGTATGTAAGAACTTGACACTATTTCTAACCA	1500

RESULT 3

US-08-525-058A-3

; Sequence 3, Application US/08525058A

; Patent No. 5770420

; GENERAL INFORMATION:

; APPLICANT: LOWE, JOHN B.

; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS

; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,

; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.

; STREET: 1755 Jefferson Davis Highway, Fourth Floor

; CITY: Arlington

; STATE: Virginia

; COUNTRY: U.S.A.

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/525,058A

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Lavalleye, Jean-Paul M. P.

; REGISTRATION NUMBER: 31,451

; REFERENCE/DOCKET NUMBER: 2363-060-55

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703)521-4500

TELEFAX: (703)486-2347  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1500 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ANTI-SENSE: NO  
US-08-525-058A-3

Query Match 100.0%; Score 1500; DB 1; Length 1500;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTTCCCTTTAGACTCTTCTTGAAGTGAAGTACCGATTCTGCTGAAGACCTCGCGCT 60  
Db 1 CCTTCCCTTTAGACTCTTCTTGAAGTGAAGTACCGATTCTGCTGAAGACCTCGCGCT 60  
Qy 61 CTCAGGCTCTGGAGTTGGAACCTGTACCTTCTTCTGCTGAGCCCTGCTCCTT 120  
Db 61 CTCAGGCTCTGGAGTTGGAACCTGTACCTTCTTCTGCTGAGCCCTGCTCCTT 120  
Qy 121 AGGAGGCCAGAGCTCGACAGAACTCGGTTGCTTGTGTTTGGTGGGGAACACAG 180  
Db 121 AGGAGGCCAGAGCTCGACAGAACTCGGTTGCTTGTGTTTGGTGGGGAACACAG 180  
Qy 181 CTGAGGATGAGGCTGACTTTTGAATCAAGAGATCTGCTTACCCAGTCTCTGGAATTA 240  
Db 181 CTGAGGATGAGGCTGACTTTTGAATCAAGAGATCTGCTTACCCAGTCTCTGGAATTA 240  
Qy 241 AGGCTGTACTACATTTGCTTGGACCTAAGATTTTCATGATCACTATGCTTCAAGATCTC 300  
Db 241 AGGCTGTACTACATTTGCTTGGACCTAAGATTTTCATGATCACTATGCTTCAAGATCTC 300  
Qy 301 CATGTCAACAAGATCTCCATGTCAGATCCAAAGTCAGAAACAGTCTTCCATCTCAAGA 360  
Db 301 CATGTCAACAAGATCTCCATGTCAGATCCAAAGTCAGAAACAGTCTTCCATCTCAAGA 360  
Qy 361 TCTGGATCACAGGAGAAATTAATGAATGTCAGGGAAGTAATCTTCTTCACTGATTT 420  
Db 361 TCTGGATCACAGGAGAAATTAATGAATGTCAGGGAAGTAATCTTCTTCACTGATTT 420  
Qy 421 GTCTCAACCGTGGTGTGCTGTTTGGGAATATGTCACAGAAATTCACAGAGTTGGTGAG 480  
Db 421 GTCTCAACCGTGGTGTGCTGTTTGGGAATATGTCACAGAAATTCACAGAGTTGGTGAG 480  
Qy 481 AACAGATGGCAGAGGACTGTGTTTCCCAAGCTGGTTTAAATGGGACCCACAGTTAT 540  
Db 481 AACAGATGGCAGAGGACTGTGTTTCCCAAGCTGGTTTAAATGGGACCCACAGTTAT 540  
Qy 541 CAAGAACAACAGTGAAGAGGAGGAGAAAGGTAGAAATGGAGATCGCATTTGAAGAG 600  
Db 541 CAAGAACAACAGTGAAGAGGAGGAGAAAGGTAGAAATGGAGATCGCATTTGAAGAG 600  
Qy 601 CCTCAGCTATGGGACTGGTTCAATCCAAAGAACCCCGGATGTTTGGACAGTACCCCG 660  
Db 601 CCTCAGCTATGGGACTGGTTCAATCCAAAGAACCCCGGATGTTTGGACAGTACCCCG 660  
Qy 661 TGGAGGCGCCGATTTGTTGGGAGGCACTTATGACACAGCTCTGCTGGAAGAGTACTAC 720  
Db 661 TGGAGGCGCCGATTTGTTGGGAGGCACTTATGACACAGCTCTGCTGGAAGAGTACTAC 720  
Qy 721 GCCACAGAAGTCACTGTGGGCTGACGTGTTTGGTGGGAAAGTACATTGAGCAT 780  
Db 721 GCCACAGAAGTCACTGTGGGCTGACGTGTTTGGTGGGAAAGTACATTGAGCAT 780  
Qy 781 TACTTAGAAGACTTTCTGGAGTCTGCTGACATGTACTTTTATGCTGTTGGCCATCGGGTCATA 840  
Db 781 TACTTAGAAGACTTTCTGGAGTCTGCTGACATGTACTTTTATGCTGTTGGCCATCGGGTCATA 840

Qy 841 TTTTACGTATGATAGACGACACCTCCCGGATGCTGTGTCACCTGAACCCCTTACAT 900  
Db 841 TTTTACGTATGATAGACGACACCTCCCGGATGCTGTGTCACCTGAACCCCTTACAT 900  
Qy 901 TCCTTACAAGTCTTTGAGATCAGTCTCAGAAAGGTTGGCAGGATATCAGCATGATCGC 960  
Db 901 TCCTTACAAGTCTTTGAGATCAGTCTCAGAAAGGTTGGCAGGATATCAGCATGATCGC 960  
Qy 961 ATGAAGACCATTTGGGAGCACATCTGCCCCACATCCACAGCAGAGTTCCTCTTTC 1020  
Db 961 ATGAAGACCATTTGGGAGCACATCTGCCCCACATCCACAGCAGAGTTCCTCTTTC 1020  
Qy 1021 TGCATGAGCTGGATCAAGTCTTTTCAAGACAATTCCTGGGGTGAACCTCTGGCCAGCTG 1080  
Db 1021 TGCATGAGCTGGATCAAGTCTTTTCAAGACAATTCCTGGGGTGAACCTCTGGCCAGCTG 1080  
Qy 1081 GTAGCACAGCTCCAGGCTGCTGTACAAAGGCCAGTCCCGAGAGTTCCACCTATGAGAGG 1140  
Db 1081 GTAGCACAGCTCCAGGCTGCTGTGTGTTACAAAGGCCAGTCCCGAGAGTTCCACCTATGAGAGG 1140  
Qy 1141 CGGGAACCTGTCGGCGCGTACATTCATTCGAGAGAGGGGATTTTACTACACCGCGCC 1200  
Db 1141 CGGGAACCTGTCGGCGCGTACATTCATTCGAGAGAGGGGATTTTACTACACCGCGCC 1200  
Qy 1201 ATTTTGGAGGAACGCTTACTACATTCCTCAACCTCACAGGAGTCTTTAAGGGATC 1260  
Db 1201 ATTTTGGAGGAACGCTTACTACATTCCTCAACCTCACAGGAGTCTTTAAGGGATC 1260  
Qy 1261 CTCAGGACAAGAATGATGATGAAGCCAGTGGCATGATGAGAGCCACTCAACAAA 1320  
Db 1261 CTCAGGACAAGAATGATGATGAAGCCAGTGGCATGATGAGAGCCACTCAACAAA 1320  
Qy 1321 TACTTCTTTTCAACAAACCCACTTCTATCTCCAGAGTATTTGCTGGGACTATCAG 1380  
Db 1321 TACTTCTTTTCAACAAACCCACTTCTATCTCCAGAGTATTTGCTGGGACTATCAG 1380  
Qy 1381 ATAGCCTGCTTCAGATATTTAAAGTCTCAAGTGTGCTGGCAGAGTATTAAT 1440  
Db 1381 ATAGCCTGCTTCAGATATTTAAAGTCTCAAGTGTGCTGGCAGAGTATTAAT 1440  
Qy 1441 TTGGTTAGAAATATGCTGACTTCAAAATTTGATGGAACCTTGACACTATTTCTAACCA 1500  
Db 1441 TTGGTTAGAAATATGCTGACTTCAAAATTTGATGGAACCTTGACACTATTTCTAACCA 1500

RESULT 4  
US-08-696-731-3  
; Sequence 3, Application US/08696731  
; Patent No. 5955347  
; GENERAL INFORMATION:  
; APPLICANT: LOWE, JOHN B.  
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
; OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/696.731  
; FILING DATE: 14-AUG-1996  
; CLASSIFICATION: 435



PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/393,246  
FILING DATE:  
APPLICANT: US 08/220,433  
FILING DATE: 30-MAR-1994  
APPLICANT: US 07/914,281  
FILING DATE: 20-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Lavalleye, Jean-Paul M. P.  
REGISTRATION NUMBER: 31,451  
REFERENCE/DOCKET NUMBER: 2363-060-55  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)521-4500  
TELEFAX: (703)486-2347  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1500 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: cDNA  
ANTI-SENSE: NO  
US-08-696-731-3

Query Match 100.0%; Score 1500; DB 2; Length 1500;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTTCCCTTTAGACTCTTCTTGAATGAGAAGTACCGATTCTGCTGAAGACCTCGCGCT 60  
DB 1 CCTTCCCTTTAGACTCTTCTTGAATGAGAAGTACCGATTCTGCTGAAGACCTCGCGCT 60

QY 61 CTGAGGCTCTGGAGTGGAACTCTGACCTCTCTTCTCTGCTGAGCCCTGCCTCTT 120  
DB 61 CTGAGGCTCTGGAGTGGAACTCTGACCTCTCTTCTCTGCTGAGCCCTGCCTCTT 120

QY 121 AGGAGCCAGAGCTCGACAGAACTCGGTTGCTTGTGCTTGTGCTTGGAGGAAACACAG 180  
DB 121 AGGAGCCAGAGCTCGACAGAACTCGGTTGCTTGTGCTTGTGCTTGGAGGAAACACAG 180

QY 181 CTGAGATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCAGTCTCCTGGAATTA 240  
DB 181 CTGAGATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCAGTCTCCTGGAATTA 240

QY 241 AGGCTCTACTACATTTGCTGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCTC 300  
DB 241 AGGCTCTACTACATTTGCTGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCTC 300

QY 301 CATGTCAACAGATCTCCATGTCAAGATCCAAGTCAAGAAACAAAGTCTTCCATCCTCAAGA 360  
DB 301 CATGTCAACAGATCTCCATGTCAAGATCCAAGTCAAGAAACAAAGTCTTCCATCCTCAAGA 360

QY 361 TCTGGATCAGAGGAAATTAATGATGTCAAGGAAAGTAAATCTCTGTTGATGCTGATT 420  
DB 361 TCTGGATCAGAGGAAATTAATGATGTCAAGGAAAGTAAATCTCTGTTGATGCTGATT 420

QY 421 GTCTCAACCTGTGCTGCTGTTTGGGAATATGTCAACAGAAATTCAGAGGTTGGTGAG 480  
DB 421 GTCTCAACCTGTGCTGCTGTTTGGGAATATGTCAACAGAAATTCAGAGGTTGGTGAG 480

QY 481 AACAGATGGCAGAGGACTGGTGTCTCCCAAGCTGGTTTAAATGGGACCCACAGTTAT 540  
DB 481 AACAGATGGCAGAGGACTGGTGTCTCCCAAGCTGGTTTAAATGGGACCCACAGTTAT 540

QY 541 CAAGAAGCAACGTAGAAGCAGGAGAGAAAGGTTAGAAATGAGATCCGATTTGAAGAG 600  
DB 541 CAAGAAGCAACGTAGAAGCAGGAGAGAAAGGTTAGAAATGAGATCCGATTTGAAGAG 600

QY 601 CCTCAGCTATGGAGTGGTTCAATCCAAAGACCGCCCGGATGTTTGCAGATGACCCCG 660  
DB 601 CCTCAGCTATGGAGTGGTTCAATCCAAAGACCGCCCGGATGTTTGCAGATGACCCCG 660

QY 661 TGAAGGCGCGGATTTGTGGGAGGACACTTATGACACAGCTCTGCTGGAAAAAGTACTAC 720  
DB 661 TGAAGGCGCGGATTTGTGGGAGGACACTTATGACACAGCTCTGCTGGAAAAAGTACTAC 720

QY 721 GCCACAGAAACTCACTGTGGGCTGACAGTGTTCCTGTGGGAAAGTACATTTGAGCAT 780  
DB 721 GCCACAGAAACTCACTGTGGGCTGACAGTGTTCCTGTGGGAAAGTACATTTGAGCAT 780

QY 781 TACTTGAAGACACTTTCTGGAGTCTGCTGACATGTACTTTCATGTTGGCCATCGGTCATA 840  
DB 781 TACTTGAAGACACTTTCTGGAGTCTGCTGACATGTACTTTCATGTTGGCCATCGGTCATA 840

QY 841 TTTTACGTCATGATAGACACACTCTCCGATGCTCTGTCGACCTGAACCCCTCTACAT 900  
DB 841 TTTTACGTCATGATAGACACACTCTCCGATGCTCTGTCGACCTGAACCCCTCTACAT 900

QY 901 TCCTTACAAGTCTTTGAGATCAGTCTGAGAAGAGTGGCAGGATATCAGCATGATCGC 960  
DB 901 TCCTTACAAGTCTTTGAGATCAGTCTGAGAAGAGTGGCAGGATATCAGCATGATCGC 960

QY 961 ATGAAGACCAATTGGGAGACACATCTCGCCACATCCAGCAGGATCGACTTCCTCTTC 1020  
DB 961 ATGAAGACCAATTGGGAGACACATCTCGCCACATCCAGCAGGATCGACTTCCTCTTC 1020

QY 1021 TGCATGACGCTGGATCAAGTCTTTCAAGACAACTTCGCGGTGGAACCTCTGGGCCAGCTG 1080  
DB 1021 TGCATGACGCTGGATCAAGTCTTTCAAGACAACTTCGCGGTGGAACCTCTGGGCCAGCTG 1080

QY 1081 GTAGCAGAGTCCAGGCTGTGTGTACAGGCGAGTCCCGAGAAAGTTCACCTATGAGAGG 1140  
DB 1081 GTAGCAGAGTCCAGGCTGTGTGTACAGGCGAGTCCCGAGAAAGTTCACCTATGAGAGG 1140

QY 1141 CGGAACTGTGCGCGGTACATTCATTCGAGAGGAGGATTTTACTACCGCGGCC 1200  
DB 1141 CGGAACTGTGCGCGGTACATTCATTCGAGAGGAGGATTTTACTACCGCGGCC 1200

QY 1201 ATTTTGGAGAAAGCGCTACTCACTTCTCAACCTCAACGAGGAGTCTTTAAGGGATC 1260  
DB 1201 ATTTTGGAGAAAGCGCTACTCACTTCTCAACCTCAACGAGGAGTCTTTAAGGGATC 1260

QY 1261 CTCAGGACAGAAACATGACATAGAGCCAGTGGCATGATGAGAGCCACCTCAACAA 1320  
DB 1261 CTCAGGACAGAAACATGACATAGAGCCAGTGGCATGATGAGAGCCACCTCAACAA 1320

QY 1321 TACTTCTTTTCAACAAACCCACTTAAATCTATCTCCAGAGTATTCCTGGGACTATCAG 1380  
DB 1321 TACTTCTTTTCAACAAACCCACTTAAATCTATCTCCAGAGTATTCCTGGGACTATCAG 1380

QY 1381 ATAGGCTGCTTTCAGATATTTAAAGTGTCAAGGTAGCTTTGGCAGACAAAGAGTATAT 1440  
DB 1381 ATAGGCTGCTTTCAGATATTTAAAGTGTCAAGGTAGCTTTGGCAGACAAAGAGTATAT 1440

QY 1441 TTGGTTAGAATAATGTCTGACTTCAAAATTTGTGATGGAACCTTGACACTATTTCTACCA 1500  
DB 1441 TTGGTTAGAATAATGTCTGACTTCAAAATTTGTGATGGAACCTTGACACTATTTCTACCA 1500

RESULT 5  
US-09-042-531-3  
Sequence 3, Application US/09042531  
Patent No. 6268193  
GENERAL INFORMATION:  
APPLICANT: LOWE, JOHN B.  
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
P.C.  
STREET: 1755 Jefferson Davis Highway, Fourth Floor

CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER: IBM PC compatible  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLYING APPLICATION NUMBER: US/09/042,531  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLYING APPLICATION NUMBER: US/08/393,246  
FILING DATE:  
APPLICATION NUMBER: US 08/220,433  
FILING DATE: 30-MAR-1994  
APPLICATION NUMBER: US 07/914,281  
FILING DATE: 20-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Lavalleye, Jean-Paul M. P.  
REGISTRATION NUMBER: 31,451  
REFERENCE/DOCKET NUMBER: 2363-060-55  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)521-4500  
TELEFAX: (703)486-2347  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1500 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: CDNA  
ANTI-SENSE: NO

US-09-042-531-3

Query Match 100.0%; Score 1500; DB 4; Length 1500;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CTCTCCCTGTAGACTCTTTGGAAATGAGAACTACCGATTCTGCTGAAGACCTCGCGCT	60
Db	1	CTCTCCCTGTAGACTCTTTGGAAATGAGAACTACCGATTCTGCTGAAGACCTCGCGCT	60
Qy	61	CTCAGGCTCTGGAGTGGAAACCTGTACCTTCCTCTGCTGAGCCCTGCTCCTT	120
Db	61	CTCAGGCTCTGGAGTGGAAACCTGTACCTTCCTCTGCTGAGCCCTGCTCCTT	120
Qy	121	AGGCAGGCCAGAGCTCGACAGAACTCGGTTGCTTTGCTTTGCTTTGGAGGGAACACAG	180
Db	121	AGGCAGGCCAGAGCTCGACAGAACTCGGTTGCTTTGCTTTGCTTTGGAGGGAACACAG	180
Qy	181	CTGACAGTAGAGCTGACCTTTGAATCAAGAGATCTGCTTACCCAGTCTCTGGAATTA	240
Db	181	CTGACAGTAGAGCTGACCTTTGAATCAAGAGATCTGCTTACCCAGTCTCTGGAATTA	240
Qy	241	AGGCTGTACTACATTTGCTTGGACCTGACCTAAGATTTCATGATCACTTCAAGATCTC	300
Db	241	AGGCTGTACTACATTTGCTTGGACCTGACCTAAGATTTCATGATCACTTCAAGATCTC	300
Qy	301	CATGTCACAAAGATCTCCATGTCGAAGTCCAAAGTCAGAAACAAAGTCTTCCATCTCAAGA	360
Db	301	CATGTCACAAAGATCTCCATGTCGAAGTCCAAAGTCAGAAACAAAGTCTTCCATCTCAAGA	360
Qy	361	TCTGGATCACAGGAGAAATTAATGAATGTCGAAGGAAAGTAATCTGCTGATGCTGATT	420
Db	361	TCTGGATCACAGGAGAAATTAATGAATGTCGAAGGAAAGTAATCTGCTGATGCTGATT	420
Qy	421	CTCTCAACCGTGGTGTGCTGCTTTGGGAATATGTCAACAGAAATCCAGAGGTTGGTGAG	480
Db	421	CTCTCAACCGTGGTGTGCTGCTTTGGGAATATGTCAACAGAAATCCAGAGGTTGGTGAG	480

...

Sequence 2, Application PC/TUS9100899  
GENERAL INFORMATION:  
APPLICANT: Lowe, John B.  
TITLE OF INVENTION: Method and Products For the Synthesis of  
TITLE OF INVENTION: Oligosaccharide Structures on Glycoproteins, Glycolipids,  
TITLE OF INVENTION: or as Free Molecules, and For the Isolation of Cloned  
TITLE OF INVENTION: Genetic Sequences That Determine These Structures  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUBAUER,  
ADDRESSEE: P.C.  
STREET: 1755 Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/00899  
FILING DATE: 19910214  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Lavalleye Ph.D., Jean-Paul  
REGISTRATION NUMBER: 31,451  
REFERENCE/DOCKET NUMBER: 2363-021-55 PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)521-5940  
TELEFAX: (703)486-2347  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1500 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: cdna  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE: Blood  
TISSUE TYPE: Blood  
PCT-US91-00899-2

Query Match 97.0%; Score 1454.6; DB 5; Length 1500;  
Best Local Similarity 98.9%; Pred. No. 0; Mismatches 0; Indels 2; Gaps 2;  
Matches 1485; Conservative 0; Mismatches 14; Indels 2; Gaps 2;

QY 1 CCTTCCCTGTAGACTCTTCTTGGAAATGAGAGTACCGATTCTGCTGAAGACCTCGCGT 60  
DB 1 CCTTCCCTGTAGACTCTTCTTGGAAATGAGAGTACCGATTCTGCTGAAGACCTCGCGT 60

QY 61 CTCAGGCTCTGGAGTGGAAACCTCTACCTTTCCTTCTGCTGAGCCCTGCTTCTT 120  
DB 61 CTCAGGCTCTGGAGTGGAAACCTCTACCTTTCCTTCTGCTGAGCCCTGCTTCTT 120

QY 121 AGCAGGCCAGACTCGACAGAACTCGGTTCCTTCTGCTTTCCTTGGAGGAAACAG 180  
DB 121 AGCAGGCCAGACTCGACAGAACTCGGTTCCTTCTGCTTTCCTTGGAGGAAACAG 180

QY 181 CTGACGATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCAGTCTCTCGAATTA 240  
DB 181 CTGACGATGAGGCTGACTTTGAACTCAAGAGATCTGCTTACCCAGTCTCTCGAATTA 240

QY 241 AGCCCTGTACTACATTTGCTTGGACCTAAGATTTTCATGATCACTATGCTTCAAGATCTC 300  
DB 241 AGCCCTGTACTACATTTGCTTGGACCTAAGATTTTCATGATCACTATGCTTCAAGATCTC 300

QY 301 CATGTCAACAGATCTCCATGTCAAGATCCAAGTCAAGAAACAGTCTTCCATCCTCAAGA 360  
DB 301 CATGTCAACAGATCTCCATGTCAAGATCCAAGTCAAGAAACAGTCTTCCATCCTCAAGA 360

QY 361 TCTGGATCACAGGAGAAAAATAATGTAAGGAAAAAGTAATCTCTGTGATGCTGATT 420  
DB 361 TCTGGATCACAGGAGAAAAATAATGTAAGGAAAAAGTAATCTCTGTGATGCTGATT 420

QY 421 GTCTCAACCGTGTGCTGCTGTTTGGGAATATGTCAACAGAAATTCACAGGTTGGTGA 480  
DB 421 GTCTCAACCGTGTGCTGCTGTTTGGGAATATGTCAACAGAAATTCACAGGTTGGTGA 480

QY 481 AACAGATGCGAGAGGACTGGTGGTTCCCAAGCTGTTTAAAAATGGAGCCACAGTTAT 540  
DB 481 AACAGATGCGAGAGGACTGGTGGTTCCCAAGCTGTTTAAAAATGGAGCCACAGTTAT 540

QY 541 CAAGAAGACAACGTAGAAGGCGGAGAGAAAGGGTAGAAATGGAGATCGCATTTGAAGAG 600  
DB 541 CAAGAAGACAACGTAGAAGGCGGAGAGAAAGGGTAGAAATGGAGATCGCATTTGAAGAG 600

QY 601 CCTCAGCTATGGGACTGTTCAATCCAAAGACCCGCCCGGATGTTTGTGACAGTGAACCCG 660  
DB 601 CCTCAGCTATGGGACTGTTCAATCCAAAGACCCGCCCGGATGTTTGTGACAGTGAACCCG 660

QY 661 TGGAGGCGCGGATGTTGGGAAGGCACTTATGACACAGCTCTGCTGGAAGAGTACTAC 720  
DB 661 TGGAGGCGCGGATGTTGGGAAGGCACTTATGACACAGCTCTGCTGGAAGAGTACTAC 720

QY 721 GCCACACAGAACTACTGTGGGGCTGACAGTGTGCTGTGGGAAAGTACATTTGAGCAT 780  
DB 721 GCCACACAGAACTACTGTGGGGCTGACAGTGTGCTGTGGGAAAGTACATTTGAGCAT 780

QY 781 TACTTGAAGACTTTCTGGAGTCTGCTGACATGATCTTCAATGTTGGCCATCGGTCATA 840  
DB 781 TACTTGAAGACTTTCTGGAGTCTGCTGACATGATCTTCAATGTTGGCCATCGGTCATA 840

QY 841 TTTTACGTCATGATACAGACACCTCCCGGATGCTGCTGACACCTGACACCTGACAT 900  
DB 841 TTTTACGTCATGATACAGACACCTCCCGGATGCTGCTGACACCTGACACCTGACAT 900

QY 901 TCCTTACAAGTCTTTGAGATCAGGTCTGAGAAGAGTGGCAGGATATCAGCATGATGCGC 960  
DB 901 TCCTTACAAGTCTTTGAGATCAGGTCTGAGAAGAGTGGCAGGATATCAGCATGATGCGC 960

QY 961 ATGAAGACCATTTGGGAGGACATCTTGGCCACATCCAGCAGGAGTGCATCTCTTTC 1020  
DB 961 ATGAAGACCATTTGGGAGGACATCTTGGCCACATCCAGCAGGAGTGCATCTCTTTC 1020

QY 1021 TGCAATGAGCTGATCAAGTCTTTCAAGACACTTCGGGGTGAAACTCTTGGGCCAGCTG 1080  
DB 1021 TGCAATGAGCTGATCAAGTCTTTCAAGACACTTCGGGGTGAAACTCTTGGGCCAGCTG 1080

QY 1081 GTAGCAGCTCCAGGCTGCTGATACAGGCGGAGTCCCGAGAGTTCACCTATGAGAGG 1140  
DB 1081 GTAGCAGCTCCAGGCTGCTGATACAGGCGGAGTCCCGAGAGTTCACCTATGAGAGG 1140

QY 1141 CGGGAATCTGCGGCCGCTGATCTTCCATTCGAGAGGAGGATTTTACTACACGCGGCC 1200  
DB 1141 CGGGAATCTGCGGCCGCTGATCTTCCATTCGAGAGGAGGATTTTACTACACGCGGCC 1200

QY 1201 ATTTTGGAGGAGGCTTACTACATTTCTCAACCTCACCAGGAGTGTCTTAAAGGGAT 1259  
DB 1201 ATTTTGGAGGAGGCTTACTACATTTCTCAACCTCACCAGGAGTGTCTTAAAGGGAT 1259

QY 1260 CTCCAGGACAAAGACATGACATAGAAGCCAGTGGCATGATGAGAGCCACCTCAACAA 1319  
DB 1260 CTCCAGGACAAAGACATGACATAGAAGCCAGTGGCATGATGAGAGCCACCTCAACAA 1319

QY 1320 ATACTTCTTTTCAACAAACCCACTAAAATCTTATCTCCAGATATTTGCTGGGACTATCA 1379  
DB 1320 ATACTTCTTTTCAACAAACCCACTAAAATCTTATCTCCAGATATTTGCTGGGACTATCA 1379

QY 1380 GATAGGCTCTGCTTCAAGATTTAAAGTGTCAAGTGTGCTTGGCAGACAAAGATATA 1439  
DB 1380 GATAGGCTCTGCTTCAAGATTTAAAGTGTCAAGTGTGCTTGGCAGACAAAGATATA 1439



QY 1284 AGAAGCCAGTGGCATGATGAGAGCCACCTCAACAAATACTTCTTTTCAACAAACCCAC 1343  
Db 1383 AGAAGCCAGTGGCATGATGAGAGCCACCTCAACAAATACTTCTTTTCAACAAACCCAC 1442  
QY 1344 TAAATCTATCTCCAGAGTATTTGCTGGGACTATCAGATAGGCGCTGCTTTTCAGATATTA 1403  
Db 1443 TAAATCTATCTCCAGAGTATTTGCTGGGACTATCAGATAGGCGCTGCTTTTCAGATATTA 1502  
QY 1404 AAGTGTCAAGCTAGCTTGGCAGACAAAGAGTATAATTTGTTAGAAATATATGCTGACT 1463  
Db 1503 AAGTGTCAAGCTAGCTTGGCAGACAAAGAGTATAATTTGTTAGAAATATATGCTGACT 1562  
QY 1464 TCAATTTGATGGAACTTTTGACACTATTTCT 1495  
Db 1563 TCAATTTGATGGAACTTTTGACACTATTTACT 1594

RESULT 8  
US-08-704-548-3  
; Sequence 3, Application US/08704548  
; Patent No. 5879675  
; GENERAL INFORMATION:  
; APPLICANT: GALILI, URI  
; TITLE OF INVENTION: REPIK, PATRICIA M.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR VACCINES  
; TITLE OF INVENTION: COMPRISING ALPHA-GALACTOSYL EPITOPES  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEIDEL, GONDA, LAVORNA & MONACO, P.C.  
; STREET: Suite 1800, Two Penn Center Plaza  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19102

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/704,548  
; FILING DATE: 11-SEP-1996  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Monaco, Daniel A.  
; REGISTRATION NUMBER: 30,480  
; REFERENCE/DOCKET NUMBER: 8760-2 CII  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-8383  
; TELEFAX: (215) 568-5549  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1131 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA

US-08-704-548-3

Query Match 67.1%; Score 1005.8; DB 2; Length 1131;  
Best Local Similarity 95.3%; Pred. No. 3.7e-280;  
Matches 1078; Conservative 0; Mismatches 2; Indels 51; Gaps 2;

QY 382 ATGAATGTCAAGGAAAGTAATCTCTTGTATGCTGATGCTCAACCGGTGGTGTGCGT 441  
Db 1 ATGAATGTCAAGGAAAGTAATCTCTTGTATGCTGATGCTCAACCGGTGGTGTGCGT 60  
QY 442 TTTTGGGAATATGTCACACG-----AATT 465  
Db 61 TTTTGGGAATATGTCACACGCGCCACGCGCTCTTCTTGTGGATATATCACACAAAAT 120

QY 466 CCAGAGTTGGTGGAGAACAGATGGCAGAGGACTGGTGGTTCCCAAGCTGTTTAAAAAT 525  
Db 121 CCAGAGTTGGTGGAGAACAGATGGCAGAGGACTGGTGGTTCCCAAGCTGTTTAAAAAT 180  
QY 526 GGGACCCACAGTTATCAAGAGACACAGTAGAAGGACGGAGAGAAAAGGTAGAAATGGA 585  
Db 181 GGGACCCACAGTTATCAAGAGAGACAGTAGAAGGACGGAGAGAAAAGGTAGAAATGGA 240  
QY 586 GATCGCATTTGA-----AGAGCCTCAGCTATGGGACTGGTTCAATCCAAAG 630  
Db 241 GATCGCATTTGAGGAAGATGACACACAGAGCCTCAGCTATGGGACTGGTTCAATCCAAAG 300  
QY 631 AACCGCCGGATGTTTGGACAGTGACCCCGTGGAAAGCGCGGATGTTGTTGGGAAGGCACT 690  
Db 301 AACCGCCGGATGTTTGGACAGTGACCCCGTGGAAAGCGCGGATGTTGTTGGGAAGGCACT 360  
QY 691 TATGACACAGCTCTGCTGGAAAGTACTAGCCACACACAGAACTCACTGTGGGCTGACA 750  
Db 361 TATGACACAGCTCTGCTGGAAAGTACTAGCCACACACAGAACTCACTGTGGGCTGACA 420  
QY 751 GTGTTTGTCTGGGAAAGTACTAGGACATTTAGAGACTTTTCTGGAGTCTGCTGAC 810  
Db 421 GTGTTTGTCTGGGAAAGTACTAGGACATTTAGAGACTTTTCTGGAGTCTGCTGAC 480  
QY 811 ATGACTTCAATGTTGGCCATCGGGTCATATTTTACGTCATATAGACACACCTCCCGG 870  
Db 481 ATGACTTCAATGTTGGCCATCGGGTCATATTTTACGTCATATAGACACACCTCCCGG 540  
QY 871 ATGCTGTCTGTCACCTGAAACCTCTACATTCCTTACAAGCTTTTGGAGATCAGGTCTGAG 930  
Db 541 ATGCTGTCTGTCACCTGAAACCTCTACATTCCTTACAAGCTTTTGGAGATCAGGTCTGAG 600  
QY 931 AAGAGTGGCAGGATATCAGCATGATGCGCATGAAGACCATTTGGGAGACACATCTCTGCC 990  
Db 601 AAGAGTGGCAGGATATCAGCATGATGCGCATGAAGACCATTTGGGAGACACATCTCTGCC 660  
QY 991 CACATCAGCAGAGGTGCGACTTCTCTTCTGTCATGAGCCTGGATCAAGTCTTTCAAGAC 1050  
Db 661 CACATCAGCAGAGGTGCGACTTCTCTTCTGTCATGAGCCTGGATCAAGTCTTTCAAGAC 720  
QY 1051 AACTTCGGGTGGAACTCTGGGCCAGCTGGTAGCAGAGCTCCAGGCTTGGTGTACAAG 1110  
Db 721 AACTTCGGGTGGAACTCTGGGCCAGCTGGTAGCAGAGCTCCAGGCTTGGTGTACAAG 780  
QY 1111 GCCAGTCCCAGAGATTTACCTATAGAGCGGGAACCTGTGGCGCGGTACATTTCCATTTC 1170  
Db 781 GCCAGTCCCAGAGATTTACCTATAGAGCGGGAACCTGTGGCGCGGTACATTTCCATTTC 840  
QY 1171 GGAGAGGGGATTTTACTACCACGGCGCATTTTGGAGGAACGCTACTCAGATTTCTC 1230  
Db 841 GGAGAGGGGATTTTACTACCACGGCGCATTTTGGAGGAACGCTACTCAGATTTCTC 900  
QY 1231 AACCTCACCAGGAGTGTCTTTAAGGGATCCTCCAGACAAAGAACATGACATAGAAGCC 1290  
Db 901 AACCTCACCAGGAGTGTCTTTAAGGGATCCTCCAGACAAAGAACATGACATAGAAGCC 960  
QY 1291 CAGTGGCATGATGAGAGCCACCTCAACAAATACTTCTTTTCAACAAACCCACTTAAATC 1350  
Db 961 CAGTGGCATGATGAGAGCCACCTCAACAAATACTTCTTTTCAACAAACCCACTTAAATC 1020  
QY 1351 CTATCTCCAGAGTATTTGCTGGGACTATCAGATAGGCGCTGCTTCAAGATATTAAGGTCTC 1410  
Db 1021 CTATCTCCAGAGTATTTGCTGGGACTATCAGATAGGCGCTGCTTCAAGATATTAAGGTCTC 1080  
QY 1411 AAGTGTGTCAGAGACAAAGAGTATAATTTGGTTAGAAATATATGCTCTCA 1461  
Db 1081 AAGTGTGTCAGAGACAAAGAGTATAATTTGGTTAGAAATATATGCTCTCA 1131

RESULT 9  
US-08-214-580A-2  
; Sequence 2, Application US/08214580A  
; Patent No. 5821117

GENERAL INFORMATION:  
APPLICANT: McKenzle, Ian F. C.  
TITLE OF INVENTION: XENOTRANSPLANTATION  
TITLE OF INVENTION: THERAPIES  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Maurice M. Klee  
STREET: 1951 Burr Street  
CITY: Fairfield  
STATE: Connecticut  
COUNTRY: USA  
ZIP: 06430  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.4 Mb storage  
COMPUTER: Dell Dimension p166v  
OPERATING SYSTEM: Windows 95  
SOFTWARE: Word for Windows 95  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/214,580A  
FILING DATE: 15-MARCH-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian Provisional  
APPLICATION NUMBER: Patent Application Serial No. 5821117 PL 7854  
FILING DATE: March 16, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Klee, Maurice M.  
REGISTRATION NUMBER: 30,399  
REFERENCE/DOCKET NUMBER: ALX-137  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (203) 255-1400  
TELEFAX: (203) 254-1101  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1423 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Double  
TOPOLOGY: Linear  
MOLECULE TYPE: cDNA to mRNA  
DESCRIPTION: galactosyl transferase,  
DESCRIPTION: full coding sequence  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Sus scrofa  
US-08-214-580A-2

Query Match 48.4%; Score: 726.6; DB 1; Length 1423;  
Best Local Similarity 78.8%; Pred. No. 1.1e-199;  
Matches 893; Conservative 0; Mismatches 234; Indels 6; Gaps 2;

Qy 354 CTCAGATCTCGATCACAGGAGAAATTAATCAATGTCAGGAGAAAGTAATCCCTGTTGAT 413  
Db 22 CCCAGCTTCTCCGATCAGGAGAAATTAATCAATGTCAGGAGAAAGTGGTCTGTCAT 81

Qy 414 GCTGATTCTCTCAACCGTGGTGTGCTGTTTGGGAATATGTCACAGAAATTCAGAGGT 473  
Db 82 GCTGCTCTCACTGCTAATGGTGTGCTGTTTGGGAATATGTCACAGAAATTCAGAGGT 141

Qy 474 TGGTGAACAGATGGCAGAGAGCTGGTGTGTTCCCAAGCTGTTTAAATGGGACCCA 533  
Db 142 TGG---CAGCAGTGTCTAGAGGGGCTGGTGTGTTCCGAGCTGGTGTAAACATGGGACTCA 198

Qy 534 CAGTTATCAACAGACACCTAGAGGACCGAGAGAAAGG---GTAGAAATGGAGATCG 590  
Db 199 CAGTTATCAACAGAGAGAGAGAGCTATAGGCAACGAAAGAAAGAAAGAGACAA 258

Qy 591 CATTGAAGAGCTCAGCTATGGGACTGGTGTCAATCCAAAGAACCGCCGGATGTTTTCAC 650  
Db 259 CAGAGGAGAGCTCCGCTAGTGGACTGGTTAATCCTCAGAAACCGCCAGAGGTCTGTCAC 318

Qy 651 AGTGACCCCGTGGAGGCGCCGATTTGTGCGGAAGGCACTTATGACACAGCTGCTGCTGA 710  
Db 319 CATAACCCAGATGGAGGCTCCAGTGGTATGGAAGGCACTTACAACAGAGCGCTCTTGA 378

Qy 711 AAAGTACTACCCACACACAGAAACTCAGTGTGGGCTGACACTGTTGCTGTGGGAAGTA 770  
Db 379 TAATATTATGCCAAACAGAAATTTACCGTGGGCTTGACGGTGTGCTGCGGAAGATA 438

Qy 771 CATTGAGCATTACTTAGAAGACTTTCTTGAGACTGCTGCTGACATGTACTTTCATGGTGGCCA 830  
Db 439 CATTGAGCATTACTTGGAGGAGTCTTAATATCTGCAATACATACTTTCATGGTGGCCA 498

Qy 831 TCGGCTCATATTTTACGTCATGATAGACGACACCTCCCGGATGCTGTCGTCACCTGAA 890  
Db 499 CAAAGTCATCTTTTACATCATGGTGGATATCTCCAGGATGCTTTCATAGAGCTGG 558

Qy 891 CCCTCTACATTTCCCTTACAAGTCTTTGAGATCAGTCTGAGAGAGAGGTGGCAGGATATCAG 950  
Db 559 TCCTCTCGGTTCTTTAAAGTGTGTTGAGATCAAGTCCGAGAGAGGTGGCAGGATATCAG 618

Qy 951 CATGATGCGCATGAAGACCATTTGGGAGCAGCATCTTGCCCCACATCCAGCAGGAGTGA 1010  
Db 619 CATGATGCGCATGAAGACCATCGGGAGCAGCATCTTGCCCCACATCCAGCAGGAGTGA 678

Qy 1011 CTTCTCTTCTGTCATGACGTCAGTCTTCAAGCAACTTTCGAGGAGTGGTGGTGGTGGT 1070  
Db 679 CTTCTCTTCTGTCATGACGTCAGTCTTCAAGCAACTTTCGAGGAGTGGTGGTGGTGGT 738

Qy 1071 GGGCCAGCTGTCAGCAGCTCGAGGCTGTCGTCAGAGGCGAGTCCCGAGAGAGTTCAC 1130  
Db 739 GGGCCAGCTGTCGTCAGCTGTCGTCAGAGGCGAGTCCCGAGAGAGTTCAC 798

Qy 1131 CTATGAGAGGCGGAACTGTCGCGCGCTACATTTCCATTCGAGAGGCGGATTTTACTA 1190  
Db 799 CTAGAGAGGCGGAGGAGTCCGAGCCTACATTTCCGTTGCGCCAGGCGGATTTTACTA 858

Qy 1191 CCACGCGCCATTTTGGAGGAGCGCTTCTCAGATCTTCAACCTCAGCAGGAGTGGT 1250  
Db 859 CCACGCGCCATTTTGGGGAACACCCACTCAGGTCTTAAACATCTCAGGAGTGGT 918

Qy 1251 TAAGGGATCTCCAGGACAAAGAACATGACATAGAGCCAGTGGCATGATGAGAGCA 1310  
Db 919 CAGGGAATCTCCAGGACAAAGAACATGACATAGAGCCAGTGGCATGATGAGAGCA 978

Qy 1311 CCTCAACAAATPACTTCTTTTCAACAAACCCACTTAAATCTCTATCCAGAGTATTGCTG 1370  
Db 979 TCTAACAAGTATTTCTTCTCAACAAACCCACTTAAATCTTATCCCCAGAAATACTGCTG 1038

Qy 1371 GGACTATCAGATAGGCTGCTTCAGATATTTAAAGTGTCAAGGTAGCTTGGCAGACAA 1430  
Db 1039 GGATTTATCATATGCGATGCTGTCGATATTTAGGATTTGTCAGAGATAGCTTGGCAGAA 1098

Qy 1431 AGACTATAATTTGTTAGAAATAATGTCGACTTCAAAATTCGATGGAACCTT 1483  
Db 1099 AGACTATAATTTGTTAGAAATAATGTCGACTTCAAAATTCGATGGAACCTT 1151

RESULT 10  
PCT-US95-07554-5  
; Sequence 5, Application PC/TUS9507554  
; GENERAL INFORMATION:  
; APPLICANT: Sandrin, Mauro S.  
; APPLICANT: Fodor, William L.  
; APPLICANT: Rother, Russell P.  
; APPLICANT: Squinto, Stephen P.  
; APPLICANT: McKenzie, Ian F. C.  
; TITLE OF INVENTION: Methods for Reducing  
; TITLE OF INVENTION: Hyperacute Rejection of Xenografts  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Maurice M. Klee  
; STREET: 1951 Burr Street  
; CITY: Fairfield

STATE: Connecticut  
COUNTRY: USA  
ZIP: 06430  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 750 kb storage  
COMPUTER: Dell 486/50  
OPERATING SYSTEM: DOS 6.2  
SOFTWARE: WordPerfect 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/07554  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/260,201  
FILING DATE: June 15, 1994  
CLASSIFICATION:  
APPLICATION NUMBER: 08/278,282  
FILING DATE: July 21, 1994  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Klee, Maurice M.  
REGISTRATION NUMBER: 30,399  
REFERENCE/DOCKET NUMBER: ALX-144.1PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (203) 255-1400  
FAX: (203) 254-1101  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1423 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Double  
TOPOLOGY: Linear  
MOLECULE TYPE: cDNA to mRNA  
DESCRIPTION: galactosyl transferase.  
DESCRIPTION: full coding sequence  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE: Sus scrofa  
PCT-US95-07554-5

Query Match 48.4%; Score 726.6; DB 5; Length 1423;  
Best Local Similarity 78.8%; Pred. No. 1.1e-199;  
Matches 893; Conservative 0; Mismatches 234; Indels 6; Gaps 2;

Qy 354 CTCAGATCTGGATCACAGGAAATAATCAATGTCACGGAAGAAAGTAACTCTGTTGAT 413  
Db 22 CCCAGCTCTGCCGATCAGGAGAAATAATGAATGTCAAAGGAAGAGTGTGTTCTGTCAT 81

Qy 414 GCTGATTTCTCAACCGTGGTGTGTTGGGAATATGTCACAGAATTCACAGAGT 473  
Db 82 GCTGTTGTTCTCACTGTAATGTTGTTTGGGATATCATCAACAGAACCAGAGT 141

Qy 474 TGGTGAGACAGATGGCAGAGGACTGGTGTCCCAAGCTGGTTTAAATAATGGGACCCA 533  
Db 142 TGG---CAGCAGTCTCAGAGGGGCTGGTGTTCAGAGCTGGTTTAAACAATGGGACTCA 198

Qy 534 CAGTTATCAGAACAACAGCTAGAAAGCAGGAGAGAAAGG---GTAGAAATGAGATCG 590  
Db 199 CAGTTACACGAAGAAGACGCTATAGGCAACGAAGAAGAAAGAAAGACAA 258

Qy 591 CATTGAAGAGCCTCAGCTATGGGACTGGTTCAATCCAAAGAACCGCCGGATGTTTGGAC 650  
Db 259 CAGAGGAGAGCTCCGCTAGTGGACTGTTTAACTCTGAGAAACGCCAGAGGTCGTGAC 318

Qy 651 AGTGACCCCGTGAAGCGCGGATGTTGTGGGAAGGCATTTATGACACAGCTGTGCTGGA 710  
Db 319 CATAAACAGATGGAAGGCTCCAGTGGTATGGGAAGGCATTTACAAACAGAGCCGCTTGA 378

Qy 711 AAGTACTACCCACACAGAACTACTGTGGGGCTCACAGTGTCTGCTGTGGGAAGTA 770  
Db 379 TAATTATTATGCCAAACAGAAATACCGTGGGCTTACCGGTTTCTGCTGTGCGAAGATA 438

Qy 771 CATTGAGCATTTACTTAGAAGACTTTCTGGAGTCTGCTGACATCTACTTCTATGCTTGGCCA 830  
Db 439 CATTGAGCATTTACTTAGAAGACTTTCTGGAGTCTGCTGACATCTACTTCTATGCTTGGCCA 498

Qy 831 TCGGTCATATTTTACGTCAATAGACAGACACTCCCGGATGCCTGCTGCTGACCTGAA 890  
Db 499 CAAAGTCATCTTTTACATCATGCTGATATCTCCAGGATGCCCTTTGATAGAGCTGG 558

Qy 891 CCCTCTACATTTCTTACAAGTCTTTGAGATCAGGTCTGAGAAGAGTGGCAGGATATCAG 950  
Db 559 TCCTCTCGCTTCTTTAAAGTGTTCAGATCAAGTCCGAGAAGAGTGGCAACACATCAG 618

Qy 951 CATGATCGGCATGAAGACCATTTGGGAGCACATCTTGGCCACATCCAGCAGAGTTCGA 1010  
Db 619 CATGATCGGCATGAAGACCATTTGGGAGCACATCTTGGCCACATCCAGCAGAGTTCGA 678

Qy 1011 CTTCCTCTCTGATGAGCGTGGATCAAGTCTTTCAAGACAACCTTCGGGGTGGAACTCT 1070  
Db 679 CTTCCTCTCTGATGAGCGTGGATCAGGTCTTTCCAAACAACCTTTGGGGTGGAGCCCT 738

Qy 1071 GGCCAGCTGCTAGCACAGCTCCAGGCTGCTGTGTACAAGGCCAGTCCCGAGAGTTTCA 1130  
Db 739 GGCCAGCTGCTGCTAGCACAGCTCCAGGCTGCTGTGTACAAGGCCAGTCCCGAGAGTTTCA 798

Qy 1131 CTATGAGAGCGGGAACATGTGGCCGCGTACATTTCCATTCGAGAGGGGATTTTACTA 1190  
Db 799 CTAGGAGCGGGAAGGAGTCCGAGCTACATTCGCTTTGGCCAGGGGATTTTATTA 858

Qy 1191 CCAGCGGCGCATTTTGGAGGAGCGCTACTACATTTCTCAACCTCCAGGAGGAGTCTT 1250  
Db 859 CCAGCGGCGCATTTTGGGGAACACCCACTCAGGTCTTAAACATCACTCAGGAGTCTT 918

Qy 1251 TAAGGGATCCTCCAGGACAAGAAACATGACATAGAGCCAGTGGCATGATGAGAGCCA 1310  
Db 919 CAAGGGAATCCTCCAGGACAAGAAATGACATAGAGCCAGTGGCATGATGAGAGCCA 978

Qy 1311 CCTCAACAAATACTTCTTTTCAACAAACCCACTTAAATCTTCTCCAGAGTATTTGCTG 1370  
Db 979 TCTAAACAAGTATTTCTTCTCAACAACCCACTTAAATCTTATCCCAAGATCTGCTG 1038

Qy 1371 GGACTATCAGATAGGCTGCTTCAGATATTTAAAGTGTCAAGTACTTGGCAGACAAA 1430  
Db 1039 GGATATCATATAGGCTGCTGCTGGATATTTAGGATTTGTCAAGATAGCTTGGCAGAAAA 1098

Qy 1431 AGAGTAAATTTGGTTAGAATAATGCTGACTTCAAAATTTGATGATGAAACTT 1483  
Db 1099 AGAGTAAATTTGGTTAGAATAATGCTGACTTAAATTTGATGATGAAACTT 1151

RESULT 11  
US-08-704-548-1  
Sequence 1 Application US/08704548  
Patent No. 5879675  
GENERAL INFORMATION:  
APPLICANT: GALILI, URI  
APPLICANT: REPIK, PATRICIA M.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR VACCINES  
TITLE OF INVENTION: COMPRISING ALPHA-CALACTOSYL EPITOPES  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.  
STREET: Suite 1800, Two Penn Center Plaza  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19102  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:



APPLICATION NUMBER: US/08/704,548  
FILING DATE: 11-SEP-1996  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Moraco, Daniel A.  
REGISTRATION NUMBER: 30,480  
REFERENCE/DOCKET NUMBER: 8760-2 CII  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-8383  
TELEFAX: (215) 568-5549  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1140 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 10..1140  
US-08-704-548-1

Query Match 46.98; Score 702.8; DB 2; Length 1140;  
Best Local Similarity 78.28; Pred. No. 7.4e-193;  
Matches 892; Conservative 0; Mismatches 197; Indels 51; Gaps 2;

QY 373 GAGAAATATGAATGTCACGAGGAAAGTAACTCTGTGTGATGCTGATGCTCAACCGTG 432  
DB 1 GAGAAATATGAATGTCACGAGGAAAGTAACTCTGTGTGATGCTGATGCTCAACCGTG 60

QY 433 GTTCTGCTGTTTTGGGAATATGTCACACAGA----- 462  
DB 61 ATTCTGTGTTTTGGGAATATATCAACAGCCACAGAGGCTCTTTCTTGTGGATATATCAC 120

QY 463 -----ATTCAGAGGTGGTGAGAACAGATGGCAGAGGACTGGTGGTCCCAAGCTGG 516  
DB 121 TCAAGAACCCAGAGGTGGTGAGAGCAGTCTCAGAGGACTGGTGGTGGTGGCTGG 180

QY 517 TTTAAATGGGACCCACAGTTATCAAGAGACACACCTAGAGGACGAGAGGAGAGAAAGGT 576  
DB 181 TTTAAATGGGACCCACAGTTATCAAGAGACACACCTAGAGGACGAGAGGAGAGAAAGGA 240

QY 577 AG-----AAATGGAGATGCGATTGAAGAGCCTCAGCTATGGGACTGGTTC 621  
DB 241 AGAGAGGAGGAACAAAAGAGGAGATGACACACAGAGCTTCGGCTATGGGACTGGTT 300

QY 622 AATCCAAGAACCCCGGATGTTTTGACAGTGACCCGCTGGAAGGCGCGATTGTGTGG 681  
DB 301 AATCCAAGAACCCCGGATGTTTTGACAGTGACCCGCTGGAAGGCGCGATTGTGTGG 360

QY 682 GAAGGCACTTATGACACAGCTCTGCTGMAAGTACTACGCGACACAGAAACTCACTGTG 741  
DB 361 GAAGGCACTTATGACACAGCTCTGCTGMAAGTACTACGCGACACAGAAACTCACTGTG 420

QY 742 GGGCTGACAGTGTGCTGTTGGGAAAGTACATGAGCATTAATGAGAGACTTTCTGGAG 801  
DB 421 GGGCTGACAGTGTGCTGTTGGGAAAGTACATGAGCATTAATGAGAGACTTTCTGGAG 480

QY 802 TCTGCTGACATGACTTTCATGCTGGCCATCGGCTCATATTTTACGTCATGATAGACGAC 861  
DB 481 TCTGCTGACATGACTTTCATGCTGGCCATCGGCTCATATTTTACGTCATGATAGACGAC 540

QY 862 ACCTCCCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 921  
DB 541 GTCTCCAAGGCGCGTTTATAGAGCTGGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600

QY 922 AGGCTGAGAGAGGTGGCAGGATATCAGCATGATGCGCATGAAGACCATTTGGGAGAC 981  
DB 601 AAGCCAGAGAGAGGTGGCAGGATATCAGCATGATGCGCATGAAGACCATTTGGGAGAC 660

QY 982 ATCTGGCCCCACATCCAGCAGGAGTGGCTTCTCTCTGTCATGGAGCTGGATCAAGTC 1041  
DB 1041 ATCTGGCCCCACATCCAGCAGGAGTGGCTTCTCTCTGTCATGGAGCTGGATCAAGTC 1041

DB 661 ATCTGGCCCCACATCCACACAGAGTTTACCTCTCTCTGTCATGGATGTGGACAGGTC 720

QY 1042 TTTCAGACAACTTTCGGGTGGAACCTCTGGCCAGCTGGTAGCACACCTCCAGGCTGG 1101

DB 721 TTCCAAGACCAATTTTGGGTAGACACCTGGCCAGCTGGTGGCTCAGCTACAGGCTGG 780

QY 1102 TGGTACAAAGCCAGTCCCGAGAGTTTACCTATGAGAGGCGGGAACCTGTCGGCCGCTGAC 1161

DB 781 TGGTACAAAGCCAGATCTCTGATGACTTTACCTATGAGAGGCGGAAAGAGTCCGCGACATAT 840

QY 1162 ATTCCATTCGAGAGGCGGATTTTACTACACAGGCGGCATTTTGGAGGAACGCTACT 1221

DB 841 ATTCCATTTGGCCAGGCGGATTTTATATACCATGACGCGATTTTGGAGGAACCGCAT 900

QY 1222 CACATTCCTCAACCTCACAGGAGTGTCTTAAAGGGGATCTCCAGGACAAACAAATGAC 1281

DB 901 CAGTTTCTCAACATCACCCAGGAGTGTCTTAAAGGAATCTCTCCGACAAAGAAATGAC 960

QY 1282 ATAGAAGCCCAAGTGGCATGATGAGAGCCACCTCAACAAATACTTCTCTTCAACAAACCC 1341

DB 961 ATAGAAGCCCAAGTGGCATGATGAGAGCCACCTCAACAAATACTTCTCTTCAACAAACCC 1020

QY 1342 ACTAAATCTCTATCTCCAGAGTATTTGCTGGGACTATCAGATAGGCTGCTTCAAGATAT 1401

DB 1021 TCTAAATCTTATCTCCAGAACTACTGCTGGGATTTATCATATAGGCTGCTTCAAGATAT 1080

QY 1402 AAAAGTGTCAAGTGTGCTGGCAGACAAAGAGATATAATTTGGTTAGAAATATGTCTGA 1461

DB 1081 AAAAGTGTCAAGTGTGCTGGCAGACAAAGAGATATAATTTGGTTAGAAATATGTCTGA 1140

RESULT 12  
US-08-621-700-1  
: Sequence 1, Application US/08621700  
: Patent No. 6153428  
: GENERAL INFORMATION:  
: APPLICANT: Gustafsson, Kenneth T.  
: Sacha, David, H.  
: TITLE OF INVENTION: '(1,3) GALACTOSYLTRANSFERASE NEGATIVE SWINE  
: NUMBER OF SEQUENCES: 24  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Carella, Bryne, Bain, Gilfillan, Cecchi, Stewart & Olstein  
: STREET: 6 Becker Farm Road  
: CITY: Roseland  
: STATE: New Jersey  
: COUNTRY: USA  
: ZIP: 07068  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: 3.5 inch diskette  
: COMPUTER: IBM PS/2  
: OPERATING SYSTEM: MS-DOS  
: SOFTWARE: ASCII  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/621,700  
: FILING DATE: 26-Mar-1996  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: 08/228,933  
: FILING DATE: <Unknown>  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Charles J. Herron  
: REGISTRATION NUMBER: 28,019  
: REFERENCE/DOCKET NUMBER: 61750-104  
: INFORMATION FOR SEQ ID NO: 1:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 1269 bases  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: HYPOTHETICAL: NO  
: ANTI-SENSE: NO  
: SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-08-621-700-1



Query Match 45.6%; Score 684.2; DB 3; Length 1269;  
Best Local Similarity 76.6%; Pred. No. 1.8e-187;  
Matches 880; Conservative 0; Mismatches 233; Indels 36; Gaps 2;

QY 371 AGGAGAAAATAATGAATGTCAGGGAAGAGTATCTCTGTGTGATGCTGATGCTCAACCG 430  
DB 5 AGGAGAAAATAATGAATGTCAGGGAAGAGTATCTCTGTGTGATGCTGATGCTCAACCG 64  
QY 431 TGGTGTGCTGTTTGGGAATATGTCACACAGAAATTCACAGAGTT----- 474  
DB 65 TAATGGTTGTGTTTGGGAATATGTCACACAGAAATTCACAGAGTT----- 474  
QY 475 -----GGTGAACAGATGGCAGAGAGGAGTGGTGGTTCCTCCCAAGCTGGT 517  
DB 125 AGTCAAAAAACCCAGAAAGTTGGCAGCAGTCTCAGAGAGGCGTGGTGTTCCTCCAGCTGGT 184  
QY 518 TTAATAATGGGACCCACAGTTATCAAGAGACACAACTAGAGAGGACGAGAGAGAAAGCG--- 574  
DB 185 TTAACAATGGGAGTACAGTTACGACGAAGAGAGACGCTATAGGCAACGAAAGGAAC 244  
QY 575 GTAGAAATGGAGATCGCATTTGAAGAGCGCTCAGCTATGGGACTGTTCAATCCAAAGAAC 634  
DB 245 AAGAAAAGACACACAGAGAGAGCTCCGCTAGTGCAGTGGTTAATCTCTGAGAAC 304  
QY 635 GCCCGATGTTTGTACAGTGACCCCGTGGAGGCGCGATTTGTGGGAGGCGACTTATG 694  
DB 305 GCCCAGAGTGTGACCATACACAGATGGAAGGCTCCAGTGGTATGGGAGGCGACTTACA 364  
QY 695 ACACAGCTCTGCGGAAAGTACTAGCCACACAGAAACTCACTGTGGGCTGACAGTGT 754  
DB 365 ACAGAGCCGCTCTAGATAATATTATGCCAAACAGAAAATACCGTGGGCTTGACCGTTT 424  
QY 755 TTCTGTGGGAAGTACATTGACGATTACTTGAAGACATTTCTGGAGTCTGCTGACATGT 814  
DB 425 TTCTGTGGGAAGTACATTGACGATTACTTGAAGACATTTCTGGAGTCTGCTGACATGT 484  
QY 815 ACTTCATGTTGGCCATCGGGTCATATTTTACGTCATGATAGACACACTCCCGGATGC 874  
DB 485 ACTTCATGTTGGCCACAAAGTCATCTTTTACATCATGTTGGTGGATGATATCTCCAGGATGC 544  
QY 875 CTGCTGTGCACCTGAACCCCTCTACATCTCTTACAAGTCTTTGAGATCAGGTCTGAGAA 934  
DB 545 CTTTGTAGAGCTGGGTCCTCTCGTTCCTTTTAAAGTGTTTGAGATCAAGTCCGAGAA 604  
QY 935 GTGGCAGGATATCAGCATGATCGCATGAAGACCATTTGGGAGACATCTCGGCCACAC 994  
DB 605 GTGGCAGACATCAGCATGATCGCATGAAGACCATCGGGAGACATCTCGGCCACAC 664  
QY 995 TCCAGCAGAGTTCGACTTCCTCTTCTGATGGACGTGGATCAAGTCTTTTCAAGACAAT 1054  
DB 665 TCCAGCAGAGTTCGACTTCCTCTTCTGATGGACGTGGATCAAGTCTTTTCAAGACAAT 724  
QY 1055 TCCGGGTGGAACTCTGGGCCACTGGTACACAGCTCCAGGCGCTGGTCAAGGCCA 1114  
DB 725 TTGGGTGGAGACCTCTGGGCCACTGGTTCAGTACAGGCTGGTGTACAGGCCAC 784  
QY 1115 GTCCCGAAGATTCACTATGAGAGCGGGAAGTGTGGCGCGGTACATTCCTCCGAG 1174  
DB 785 ATCTGACGAGTTCACTACGAGAGCGGGAAGAGTCCCGAGGCTACATTCCTTTGGCC 844  
QY 1175 AGGGGATTTTACTACGAGCGGCCATTTTGGAGGAAGCGCTACTCTCAATTTCAACC 1234  
DB 845 AGGGGATTTTATACGAGCGGCCATTTTGGGGAACACCCACTCAGGTTCTAAACA 904  
QY 1235 TCACCGAGGAGTCTTTAAGGGATCTCTCAGGACAGAAACATGACATAGAGCCCGAGT 1294  
DB 905 TCACCTCAGGAGTCTTTCAAGGGAATCTCTCAGGACAGGAAATGACATAGAGCCCGAGT 964  
QY 1295 GGCATGATGAGCCACCTCAACAAATACCTCTTTTCAACAAACCCACTTAAATCCTAT 1354  
DB 965 GGCATGATGAGCCCATCTAACAAGATATTTCTCTTCAACAAACCCACTTAAATCCTAT 1024

QY 1355 CTCCAGAGATATTCTCTGGGACTATCAGATAGGCTGCTTTCAGATATTAAGAGTGTCAAGG 1414  
DB 1025 CCCCAAGTACTGCTGGGATTTATCATATAGGCATGCTCTGGGATATTAGGATTTGTCAAG 1084  
QY 1415 TAGCTTGGCAGACAAAGAGTATAATTTGGTTAGAATAATGCTCTGACTTCAANTTTGTA 1474  
DB 1085 TAGCTTGGCAGAAAAGAGTATAATTTGGTTAGAATAATGCTCTGACTTCAANTTTGTA 1144  
QY 1475 TGGAAACTT 1483  
DB 1145 CAGCAGTTT 1153

RESULT 13  
PCT-US95-03940-1  
; Sequence 1, Application PC/TUS9503940  
; GENERAL INFORMATION:  
; APPLICANT: BIOTRANSPLANT, INC.  
; APPLICANT: THE GENERAL HOSPITAL CORPORATION  
; APPLICANT: THE UNIVERSITY OF LONDON  
; APPLICANT: BAETSCHER, Manfred W.  
; APPLICANT: GUSTAFSSON, Kenth T.  
; APPLICANT: SACHS, David H.  
; TITLE OF INVENTION: (1,3) GALACTOSYLTRANSFERASE NEGATIVE SWINE  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,  
; ADDRESSEE: Stewart & Olstein  
; STREET: 6 Becker Farm Road  
; CITY: Roseland  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07068  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch diskette  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/03940  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/228,933  
; FILING DATE: April 13, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Charles J. Herron  
; REGISTRATION NUMBER: 28,019  
; REFERENCE/DOCKET NUMBER: 61750-CIP  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1269 bases  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; HYPOTHEetical: NO  
; ANTI-SENSE: NO  
PCT-US95-03940-1

Query Match 45.6%; Score 684.2; DB 5; Length 1269;  
Best Local Similarity 76.6%; Pred. No. 1.8e-187;  
Matches 880; Conservative 0; Mismatches 233; Indels 36; Gaps 2;

QY 371 AGGAGAAAATAATGAATGTCAGGGAAGAGTATCTCTGTGTGATGCTGATGCTCAACCG 430  
DB 5 AGGAGAAAATAATGAATGTCAGGGAAGAGTATCTCTGTGTGATGCTGATGCTCAACCG 64  
QY 431 TGGTGTGCTGTTTGGGAATATGTCACACAGAAATTCACAGAGTT----- 474  
DB 65 TAATGGTTGTGTTTGGGAATATGTCACACAGAAATTCACAGAGTT----- 474  
QY 475 -----GGTGAACAGATGGCAGAGGAGTGGTGGTTCCTCCCAAGCTGGT 517

```
Db 125 AGTCAAAAAAAGTGGCAGCAGTGTCTCAGAGGGGCTGGTGGTTCCTCCAGCTGGT 184
Qy 518 TTAATAATGGGACCCACAGTTATCAACAAGACACAGCTAGAGGAGCGGAGAGAAAGG- 574
Db 185 TTAACAATGGGACTCAGCTTACCACGAGAGAGACGCTATAGGCAAGCAAAAGAAC 244
Qy 575 GTAGAAATGGAGATCGCATTTGAAGAGCCTCAGCTATGGGACTGGTTCAATCCAAAGAAC 634
Db 245 AAAGAAAGAAGACACAGAGGAGAGCTCCGCTAGTGGACTGGTTTAATCCTGAGAAC 304
Qy 635 GCCGGATGTTTTCAGAGTGAACCCGCTGGAAGGCGCGATGTTGTTGGGAAGCACTTATG 694
Db 305 GCCCAGAGGTCGTGACCATTAACAGATGGAAGGCTCCAGTGGTATGGGAAGGCACTTACA 364
Qy 695 ACACAGCTCTGCTGGAAGTACTACCCACACAGAAACTCACTGTGGGCTGACAGCTGT 754
Db 365 ACAGAGCGCTCTAGATTAATTTATGCCAACAAGAAATTAACGCTGGGCTTGACGGTTT 424
Qy 755 TTGCTGTGGGAAAGTACATTTGAGCATTTACTTAGAAGACTTTCTGGAGTCTGCTGACATGT 814
Db 425 TTGCTGTGCGAAGATACATTTGAGCATTTACTTTGGAGGAGTCTTAATATCTGCAAAATACAT 484
Qy 815 ACTTCATGTTGGCCATCGGGTGCATATTTTACGTCATGATAGAGGACACCTCCCGGATGC 874
Db 485 ACTTCATGTTGGCCACAAAAGTCATCTTTTACATCATGTTGGATGATATCTCCAGGATGC 544
Qy 875 CTGCTGTGCACTGCAACCTCTACATTTCTTACAAAGTCTTTGAGATCAGCTCTGAGAGA 934
Db 545 CTTGATAGAGCTGGGTCTCTGGGCTCTTAAAGTCTTTGAGATCAAGTCCGAGAGA 604
Qy 935 GGTGGCAGGATATCAGCATGATGCGATGAAGACCAATTTGGGAGCAGCATCTCGGCCACA 994
Db 605 GGTGGCAAGACATCAGCATGATGCGATGAAGACCAATCGGGAGCAGCATCTCGGCCACA 664
Qy 995 TCNAGCAGAGGTCGACTTCTCTTCTGTCATGACGTGGATCAAGTCTTTCAAGACAACT 1054
Db 665 TCCAGCAGAGGTTGGATCTCTCTCTGTCATGACGTGGATCAAGTCTTTCAAAACAACT 724
Qy 1055 TCGGGTGGAAACTCTCGGCGAGCTGTAGCAGCAGCTTCCAGGCTGTGTGTACAAAGCCA 1114
Db 725 TTGGGTGGAGACCTCGGCGAGCTCGGTGCTCAGCTACAGGCTGTGTGTACAAAGCAC 784
Qy 1115 GTCCCGAAGATTACCTATGAGAGCGGGAAGTGTGCGCGCGGTACATTTCCATTCGGAG 1174
Db 785 ATCCTGACGAGTTTCACTACGAGAGCGGGAAGGAGTCCGCGAGCTACATTTCCGTTTGGCC 844
Qy 1175 AGGGGATTTTACTACCGCGGCCATTTTGGAGGAGCGCCTACTACATTTCTCAACC 1234
Db 845 AGGGGATTTTATACCGCAGCCATTTTGGGGAACACCCACTCAGGTTCTTAAACA 904
Qy 1235 TCACGAGGAGTGTCTTTAAGGGATCTCCAGGACAAAGAACATGACATAGAAGCCAGT 1294
Db 905 TCACGAGGAGTGTCTTTAAGGGATCTCCAGGACAAAGAACATGACATAGAAGCCAGT 964
Qy 1295 GGCATGATGAGAGCCACTCAACAATATCTTCTTTTCAACAACCCACTAAAATCCTAT 1354
Db 965 GGCATGATGAGAGCCACTCAACAATATCTTCTTTTCAACAACCCACTAAAATCCTAT 1024
Qy 1355 CTCAGAGTATTTGCTGGGACTATCAGATAGGCTGCTCTCAGATATTTAAAGTGTCAAGG 1414
Db 1025 CCCAGAAATCTGCTGGGATATCATATAGGCAATGCTGTGGATATTTAGGATTTGTCAAGA 1084
Qy 1415 TAGCTTGGCAGAAAAAGATATAATTTGGTTAGAAATATGCTGACTTCAAAATCTGTA 1474
Db 1085 TAGCTTGGCAGAAAAAGATATAATTTGGTTAGAAATATGCTGACTTCAAAATCTGTA 1144
Qy 1475 TGGAACTT 1483
Db 1145 CAGCAGTTT 1153
```

RESULT 14

```
US-08-378-617A-7
; Sequence 7, Application US/08378617A
; Patent No. 5849991
; GENERAL INFORMATION:
; APPLICANT: d'Apice, Anthony J.F.
; APPLICANT: Pearse, Martin J.
; APPLICANT: Robins, Allan J.
; APPLICANT: Crawford, Robert J.
; APPLICANT: Rachen, Peter D.
; TITLE OF INVENTION: MATERIALS AND METHODS FOR MANAGEMENT OF
; TITLE OF INVENTION: HYPERACUTE REJECTION IN HUMAN XENOTRANSPLANTATION
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 120 South Sixth Street, Suite 2500
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30B
; APPLICATION NUMBER: US/08/378,617A
; FILING DATE: 26-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 06868/005001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (612) 335-5070
; TELEFAX: (612) 288-9696
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1412 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-378-617A-7
```

```
Query Match 45.6%; Score 684.2; DB 2; Length 1412;
Best Local Similarity 76.6%; Pred. No. 1.9e-187;
Matches 880; Conservative 0; Mismatches 233; Indels 36; Gaps 2;

Qy 371 AGGAGAAAATAATGAATGTCAAGGGAAAAGTAATCTCTGTGTGATGCTGATTTCTCAACCG 430
Db 80 AGGAGAAAATAATGAATGTCAAGGGAAAAGTGGTTCCTGTCATATGCTGCTTCAACTG 139
Qy 431 TGGTTGCTGTTTGGGAATATGTCAACAGAAATTCACAGAGTT----- 474
Db 140 TAATGGTGTGTTTGGGAATATCATCAACAGCCCGCAGAGGTTCTTTCTTCTGGATATACC 199
Qy 475 -----GCTGAGAACACAGATGGCAGAGGACTGCTGTTCCCAAGCTGGT 517
Db 200 AGTCAAAAACCCACAGAGTTGGCAGCAGTCTCAGAGGGGCTGGTGTTCGAGCTGGT 259
Qy 518 TTAATAATGGGACCCACAGTTATCAAGAAGACACACGTAGAGGACGAGAGAGAAAGG--- 574
Db 260 TTAACAATGGGACTCAGCTTACCACCAAGAAAGACGCTATAGGCAAGCAAAAGGAAC 319
Qy 575 GTAGAAATGGAGATCGCATTTGAAGAGCCTCAGCTATCGGACTGCTTCAATCCAAAGAAC 634
Db 320 AAAGAAAGAAGACACAGAGGAGAGCTTCGCTAGTGGACTGGTTTAAATCCTGAGAAAC 379
Qy 635 GCCCGATGTTTTCAGAGTGAACCCGCTGGAAGGCGCGATTTGTGTGGCAAGGCACTTATG 694
Db 380 GCCCAGAGGTCGTGACCATTAACAGATGGAAGGCTCCAGTGGTATGGGAAGGCACTTACA 439
```

[illegible]

RESULT 15

RESULT IS  
 US-08-378-617A-8  
 : Sequence 8, Application US/08378617A  
 : Patent No. 5849991  
 :  
 : GENERAL INFORMATION:  
 :  
 : APPLICANT: d'Aplice, Anthony J.F.  
 : APPLICANT: Pearse, Martin J.  
 : APPLICANT: Robins, Allan J.  
 : APPLICANT: Crawford, Robert J.  
 : APPLICANT: Rathjen, Peter D.  
 : TITLE OF INVENTION: MATERIALS AND METHODS FOR MANAGEMENT OF  
 : TITLE OF INVENTION: HYPERACUTE REJECTION IN HUMAN XENOTRANSPLANTATION  
 :  
 : NUMBER OF SEQUENCES: 33  
 :  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Fish & Richardson  
 : STREET: 120 South Sixth Street, Suite 2500  
 :

CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, version #1.30B  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/378,617A  
FILING DATE: 26-JAN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Ellinger, Mark S.  
REGISTRATION NUMBER: 34,812  
REFERENCE/DOCKET NUMBER: 06868/005001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (612) 335-5070  
TELEFAX: (612) 288-9696  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1828 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-378-617A-8

	Query Match	42.8%;	Score 642.4;	DB 2;	Length 1828;
	Best Local Similarity	75.1%;	Pred. No. 2.4e-175;		
	Matches 849;	Conservative 0;	Mismatches 236;	Indels 45;	Gaps
Qy	371	AGGAGAAATATGAATGTCAGGGGAAAAGTAACTCTGTGATGCTGATGCTCTCAACCG	430		
Db	458	AGGAGAAATATGAATGTCAGGGGAAAAGTAACTCTGTGATGCTGATGCTCTCAACCTG	517		
Qy	431	TGCTGTGCTGTTTGGGAATATGTCAC	459		
Db	518	TCATTGTGTGTTTGGGAATATATCCACAGCCCAAGAGGCTCTTTGTTCTGGATAAAC	577		
Qy	460	-----AGAATTCAGAGGTTGGTGAGACAGATGCCAGAGGACTGGTGTCCCAAGCT	514		
Db	578	CATCAAGAAACCCCAAGATTTGGTGGCAGCAGCATTCAGAAAGGCTGGTGGCTCCGAGAT	637		
Qy	515	GGTTTAAAAATGGGACCCACAGTTATCAAGAAGACAAAGCTAGAGGACGCGAGAGAAAAGG	574		
Db	638	GGTTTACAAATG-----GTTACCATGAAGAAGATGGAGACATAAACCAAGANAAGG	688		
Qy	575	GTAGAAATGGAGATCGGATTTGAAGAGCTTCAGCTATGGGACTTGGTTCATTCAAAGAACCC	634		
Db	689	AACAAGAAACGAAGACGAAGCAAGCTTTAAGCTATCGGACTTGGTTCACCCCATTTAAAC	748		
Qy	635	GCCCGGAGTTTTCAGACAGTACCCCGCTGGAAGCGCGGATTTGTTGGGAAGCGACTTTATG	694		
Db	749	GCCCGAGGTTGTGACCATGTGCAAGTGGGAAGGCTCCAGTGGTGTGGGAAGGCACTTACA	808		
Qy	695	ACACAGCTCTGCTGGAAAAGTACTACGCCACACAGAAAACCTCACTGTGGGGCTGCACGTGT	754		
Db	809	ACAGAGCGCTTTAGACAATTTATATGCCAAGCAGAAAAATACCGTCGCGCTGACGGTTT	868		
Qy	755	TTGCTGTGGGAAAGTACATTTGAGCATTTACTTTAGAAGACTTTCTGGAGTCTGCTCACATGT	814		
Db	869	TCGCGCTCGGAAGATACATTTGAGCATTTACTTTGGAGGAGTCTTTAACGTCCTCTAATANGC	928		
Qy	815	ACTTCATGTTTGCCATCGGGTGCATATTTTTCGTCATGATAGACGACACACCTCCCGGATGC	874		
Db	929	ACTTCATGTTGGGCCACCCACGTCATCTTTTATATCATGCTAGATGATCTCTCCAGGATGC	988		
Qy	875	CTGTCGTGCACCTGAACCCCTCTACATTCCTTTACAAGTCTTTGAGATCAGGCTCTCAGAGA	934		
Db	989	CTTTGATAGAGTTTGGTCTCTCGCTCTCTCAAACTGTTTAAAGATCAAGCTCTGAGA	1048		



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 17:56:14 ; Search time 252.09 Seconds  
(without alignments)  
10216.080 Million cell updates/sec

Title: US-09-863-475a-3  
Perfect score: 1500  
Sequence: 1 CCTTCCCTGTAGACTTTC.....CTTGACACTATTCTTAACCA 1500

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues  
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :				N_Geneseq_032802.*	
1:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*	2:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*	3:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*	5:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*	6:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*	8:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*	9:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*	11:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*	12:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*	14:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*	15:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*	17:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*	18:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*	20:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*	21:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*	23:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*	24:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	1500	100.0	1500	12	AAQ13331
2	1500	100.0	1500	15	AAQ56907
3	1500	100.0	1500	18	AAAT61676
4	1500	100.0	1500	24	AAAD26484
5	1434.8	95.7	3537	22	AAQ505315
6	1030.8	68.7	1116	16	AAQ04523
7	726.6	48.4	1423	15	AAQ74712
8	726.6	48.4	1423	17	AAQ12242
9	725.2	48.3	3240	22	AAQ505314

10	706	47.1	1092	19	AAV49454	Porcine alpha-1,3-
11	693.8	46.3	1131	16	AAQ04522	Marmoset alpha-1,3
12	693.8	46.3	1131	24	AAQ26482	Marmoset alpha (1,
13	684.2	45.6	1269	24	AAQ26483	Pig alpha (1, 3) g
14	681	45.4	1269	16	AAQ02892	Porcine alpha (1,3
15	677.8	45.2	1410	16	AAQ93077	Pig alpha-1,3-gala
16	664.6	44.3	1128	19	AAV49453	Porcine alpha-1,3-
17	645.6	43.0	1496	24	AAQ26480	sheep alpha (1, 3)
18	642.4	42.8	1617	24	AAQ26481	Bovine alpha (1, 3
19	632.8	42.2	1065	19	AAV49455	Porcine alpha-1,3-
20	617.2	41.1	1029	19	AAV49456	Porcine alpha-1,3-
21	583.2	38.9	1353	15	AAQ74711	Galactosyl transfe
22	561.8	37.5	1885	24	AAQ26485	Human alpha (1, 3)
23	518.2	34.5	2964	22	AAQ505303	Pig alpha-1,3 gala
24	516.2	34.4	3322	22	AAQ505322	Human alpha-1,3 ga
25	511.2	34.1	3745	22	AAQ505323	Human alpha-1,3 ga
26	479	31.9	3135	22	AAQ505316	Human alpha-1,3 ga
27	479	31.9	3558	22	AAQ505317	Human alpha-1,3 ga
28	477.4	31.8	2303	22	AAQ505332	Rhesus monkey alph
29	476.6	31.8	2848	22	AAQ505331	Rhesus monkey alph
30	475.6	31.7	2630	22	AAQ505333	Human Huma exon 6
31	241.6	16.1	826	20	AAQ65174	Human drug metabol
32	241.6	16.1	1593	22	AAQ09959	Histo-blood gp. A
33	240	16.0	1062	12	AAQ11131	Human A transferas
34	240	16.0	1062	15	AAQ68806	Baboon A allele DN
35	233.6	15.6	826	20	AAQ65175	Baboon A allele DN
36	232	15.5	826	20	AAQ65176	Baboon B allele DN
37	232	15.5	826	20	AAQ65177	Human polynucleoti
38	211	14.1	1684	22	AAQ58407	Partial sequence o
39	188.6	12.6	585	15	AAQ68826	Mouse alpha-1,3 ga
40	184.4	12.3	900	22	AAQ505308	Human polynucleoti
41	184.2	12.3	1926	22	AAQ160193	Human full-length
42	182.6	12.2	1926	22	AAQ94349	Human full-length
43	164.4	11.0	2718	22	AAQ94349	Mouse alpha-1,3 ga
44	141.8	9.5	608	22	AAQ505313	Mouse alpha-1,3 ga
45	110	7.3	1207	22	AAQ505307	Mouse alpha-1,3 ga

ALIGNMENTS

RESULT 1	AAQ13331	standard; cDNA; 1500 BP.
ID	AAQ13331	
XX	AAQ13331	
AC	AAQ13331	
DT	07-NOV-1991	(first entry)
XX	UDP-Gal:beta-D-Gal(1,4)-D-GlcNAc alpha(1,3)-	galactosyltransferase gene.
DE	DE	Glycosyltransferase; ss.
DE	DE	Glycosyltransferase; ss.
KW	XX	Mus musculus.
XX	OS	Mus musculus.
XX	XX	Key
FH	XX	Location/Qualifiers
CD	XX	277..1185
FT	XX	/*tag= a
XX	XX	WO9112340-A.
PN	XX	22-AUG-1991.
XX	XX	14-FEB-1991;
PF	XX	91WO-US00899.
XX	XX	12-DEC-1990;
PR	XX	90US-0627621.
PR	XX	14-FEB-1990;
PR	XX	90US-0479858.
XX	XX	14-FEB-1990;
XX	XX	90US-0480133.
PA	XX	(UNMI ) UNIV OF MICHIGAN.
XX	XX	Lowe JB;
PI	XX	

XX WPI; 1991-267151/36.  
DR P-PSDB; AAR13750.  
XX  
PT Isolation of gene conveying post-translational characteristic -  
PT e.g. the presence of soluble or membrane bound oligo or  
PT polysaccharide or glycosyltransferase.  
XX  
PS Claim 11; Fig 2: 155pp; English.  
XX  
CC The DNA encodes a protein sequence capable of functioning as a  
CC UDP-Gal: [beta-D-Gal(1,4)]-D-olivnac alpha(1,3)-galactosyltransferase  
CC and was isolated from pCDM7-alphaHgr. The DNA sequence can be used  
CC to construct animal cell lines with respect to posttranslational  
CC modification of the oligosaccharides on cell-surface, intracellular  
CC or secreted proteins or lipids by sub-terminal alpha(1,3) and alpha  
CC (1,4) fucose residues that represent the products of this enzyme.  
CC These can be used for the prodn. of diagnostics and therapeutics.  
CC See also AAQ13330-Q13333.  
XX  
SQ Sequence 1500 BP; 406 A; 349 C; 374 G; 371 T; 0 other;

Query Match 100.0%; Score 1500; DB 12; Length 1500;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTTCCCTTGTAGACTCTTCTTGAATGAGAACTACCGATTCTGCTGAAGACCTCGCGCT 60  
DB 1 ccttccctttagactcttcttggaaagagaaagtaacgattctgtgaagacctcgcgct 60  
QY 61 CTCAGGCTCTGGGAGTTGGAACTGACCTGACCTTCTTCTCTGCTGAGCCCTGCTCCTT 120  
DB 61 ctcaggctctgggagttggaaacctgtacotctcttctctgtgagccctgcctcctt 120  
QY 121 AGGAGGCCAGAGCTCGACAGAACTCGTTGCTTGTGCTTGTGCTTGTGGAGGAACACAG 180  
DB 121 aggcagccagagctcgacagaaactcgtgttctgttctgttcttggagggaacacag 180  
QY 181 CTGACGATGAGGCYGACTTTGAATCAAGAGATCTGCTTACCCAGTCTCTCGGAATTA 240  
DB 181 ctgacgatgagcgtacttgaactcaagagatctgtcttaccctcagctctcctggaattaa 240  
QY 241 AGGCCTGTACTACATTTGCTCGGACCTTAAGATTTTCATGATGACTATGCTTCAAGATCTC 300  
DB 241 aggcctgtactacatttgcctggacctaaagatttctcatgactatgcttcaagatctc 300  
QY 301 CATCTCAACAGATCTCCATGCTCAAGATCCAAGTCAAGAACAACTCTTCCATCTCAAGA 360  
DB 301 catgtcaacaaga tctcca tgcagaatcccaag tcaagaacaagcttctccatccctcaaga 360  
QY 361 TCTGGATCACAGGAGAAATTAATGAATGTCGAAGGAAAGTAATCTCTGTTGATGCTGATT 420  
DB 361 tctggatcacaggagaaaaataatgaatgtcaaggaaaaagtaacctgttgatgctgatt 420  
QY 421 GTCTCAACCGTGTGCTGCTTTTGGGAATATGTCACAGAAATCCAGAGTTGGTGGAG 480  
DB 421 gtctcaacggtgtgtcgtgttcttgggaatgtgtcaacagaattccagaggttggtag 480  
QY 481 AACAGATGGCAGAGGACTGCTGCTTCCAGCTGTTTAAATGGACCCACAGATTAT 540  
DB 481 aacagatggcagagagactgctgcttccagctggttttaaatggacccacagattat 540  
QY 541 CAAGAACAACAGCTAGAGGACGAGAGAAAGGTTAGAAATGAGATCGCATTTGAAGAG 600  
DB 541 caagaagaacaagctagaagagcagagagaaaaaggtagaataaggagatcgcattgaagag 600  
QY 601 CCTCAGCTATGGGACTGGTTCAATCCAAAGAACCCCGGATGTTTTCACAGTGAACCCCG 660  
DB 601 cctcagctatgggactggttcaatccaaagaaacccgcgagatgtttgacagtgaccccg 660  
QY 661 TGGAAAGGCCGATTTGTGGGAGGCACTTATCACACAGCTCTGCTGGAAAGTACTAC 720  
DB 661 tggaaagccgatttgtgggaggcacttattcacacagctctgctggaaagtactac

DB 661 tggaaagcgccgattgtgtgtggaaaggcaacttatgacacagctctgtcgtggaanaagtaactac 720  
QY 721 GCCACACAGAACTCACTGTGGGCTGACAGTGTGCTGGGAAAGTACATTTGAGCAT 780  
DB 721 gccaacacagaactcactgtgggctgacagtgcttgcgtgtgggaagatgacatlgagcat 780  
QY 781 TACTTAGAAGACTTTTCTGGAGTCTGCTGACATGTACTTTCATGTGTTGGCCATCGGGTTCATA 840  
DB 781 tacttagaagacttcttggagctctgtgacatgtacttcaatggttggccatcggggtcata 840  
QY 841 TTTTACCTCATGATAGACGACACCTCCCGATGCTGCTGCGACCTGAACCCCTCTACAT 900  
DB 841 ttttacctcatgatagacgacacactcccgatgctctgtcgtgacactgaaacctctacat 900  
QY 901 TCCTTACAAGTCTTTGAGATCAGGTCTGAGAAGAGGTGGCAGGATATCAGCATGATGGCC 960  
DB 901 tccittacaagcttcttgagatcaggtctcgagaagaggtggcaggatatacagcatgatcgcc 960  
QY 961 ATGAAGACCATTTGGGAGCAGCATCTTGGCCCCACATTCAGACACGAGGTGCGATTCTCTTC 1020  
DB 961 atgaagaccatttgggagcacatctgtgccacatccagcacgaggtcgacttctctctc 1020  
QY 1021 TGCATGCGAGCTGGATCAAGTCTTTCAAGACAACTTCCGGGTGGAACCTCTCGGCCACAGCTG 1080  
DB 1021 tgcattgacgtggagctcaagctcttcaagacaaactctcggttggaaactctctggccagctg 1080  
QY 1081 GTAGCACAGCTTCCAGCGCTGCTGTACAAAGGCCAGTCCCGAGAAGTTTCACTATGACAGG 1140  
DB 1081 gtacacagctccagcgtgtgtgtacaaagccagtcgccgagaagttcacctatlgagagg 1140  
QY 1141 CGGAACTGTGCGCGCGGTACATTTCCATTCGGAGAGGGGGATTTTACTACCACCGCGCC 1200  
DB 1141 cgggaactgttcggcgcgtacattccattcggagagggggttttactaccacgcgcgccc 1200  
QY 1201 ATTTTTCGAGGAACGCGCTACTACATTTCAACCTCAACCTCAACGAGGTGCTTTAAGGGATC 1260  
DB 1201 atttctggagaaacgctcactacatctcctaacctccacgggggtcttcaaaggagatc 1260  
QY 1261 CTCCAGGACAGAACAATGACATAGAAGCCAGTGGCATGATGAGAGCCACTCAACAAA 1320  
DB 1261 ctccaggacaagaacatgacatagaagccagtggtcatgtatgagagccactcaacaaa 1320  
QY 1321 TACTTCTTTTCAACAACCCACTTAAATCCTATCTCCAGAGTATTGCTGGGACTATCAG 1380  
DB 1321 tacttcttttcaacaacccactaaatccta tctccagagtatttctgtgggactatcag 1380  
QY 1381 ATAGCGCTTGCCTTCAGATATTAAAAGTCTCAAGGTAGCTTGGCAGACAAAAGAGTATAAT 1440  
DB 1381 ataggcctgcttcagatattaaaagtgtcaagtgctgagtagcttggcagacaaaagatataat 1440  
QY 1441 TTGTTTAGAATAATGCTGACTTCAAAATTGCTGATGGAACCTTGACACTATTTCCTAACCA 1500  
DB 1441 ttggttagaataatgtctgacttcaaatgtgtgtagaaacttgacactatttcttaacca 1500

## RESULT 2

AAQ56907  
ID AAQ56907 standard; DNA; 1500 BP.

XX  
AC AAQ56907;

XX  
DT 26-JUL-1994 (first entry)

XX  
DE DNA encoding glycosyltransferase.

XX  
KW Glycosyltransferase; fucosyltransferase; GDP-Fuc; in vitro; cell;  
KW surface; oligosaccharide; ss.

XX  
OS Homo sapiens.

XX  
FH Key Location/Qualifiers  
FT CDS 276..1474

FT /\*Cag= a



XX	Alpha(1,3)-galactosyltransferase; glycosylation; oligosaccharide:
KW	ss.
XX	
OS	Mus sp.
XX	
XX	Key Location/Qualifiers
FT	CDS 277..1461
FT	/*tag= a
XX	
PN	WO9709421-A1.
XX	
PD	13-MAR-1997.
XX	
PF	06-SEP-1996; 96WO-US13816.
XX	
PR	08-SEP-1995; 95US-0525058.
XX	
PA	(UNMI ) UNIV MICHIGAN.
PI	Legault DJ, Lowe JB;
XX	
DR	WPI; 1997-192897/17.
XX	
XX	P-PSDB; AAW13639.
PT	New recombinant fucosyltransferase proteins - useful for modifying
PT	cell surface oligosaccharide structures
XX	
PS	Example 2; Page 271-272; 329pp; English.
XX	
CC	A cDNA clone (AAT61676) codes for an enzyme (AAW13639) that functions
CC	as a UDP-Gal:beta-D-Gal(1,4)-D-GlcNAc alpha(1,3)-
CC	galactosyltransferase. It was obtd. by transfecting COS-1 cells
CC	with cDNA derived from mouse F9 teratocarcinoma cells, and
CC	screening the transfected cells for surface-localised
CC	Gal(alpha1-3)Gal linkages. The cDNA can be used to construct
CC	animal cell lines with specific capabilities with respect to post-
CC	translational modification of the oligosaccharides of expressed
CC	proteins or lipids, or to produce recombinant enzyme for use in
CC	oligosaccharide prodn.
XX	
SO	Sequence 1500 BP; 406 A; 349 C; 374 G; 371 T; 0 other;
Query Match 100.0%; Score 1500; DB 18; Length 1500;	
Best Local Similarity 100.0%; Pred. No. 0;	
Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps	
Oy	1 CCTTCCCTTGTAGACTCTTCTTGGAAATGAGAACTACCGATTCTGCTGAAGACCTCGCGCT
Db	
Oy	1 ccttcctttagactcttcttggaaatgagaagaccgattctgctgagaagacctcgcgct
Db	
Oy	61 CTAGGCTCTGGAGTTGGAAACCTGTACCTTCTTCTCTGCTGAGCCCTGCTCTCTT
Db	
Oy	61 ctgagctctggagttggaacctgtacctcttctctgctgctgagcctgctcctt
Db	
Oy	121 AGCAGGCCAGACTCGACAGAACTCGTTTGTCTTGTCTTGGAGGGAACACAG
Db	
Oy	121 agcagggccagagctcgagaaactcggttgctctgcttcttggagggaaacacag
Db	
Oy	181 CTGACGTAGGGCTGACTTTTGAACCTCAAGAGATCTCTTACCCCACTCTCTCGAATTA
Db	
Oy	181 ctgacgtagagctgacttttgaactcaagagatctgttaccttaccttgaatbaa
Db	
Oy	241 AGGCTGTACTACATTTGGCTTGACACCTAGATTTTCATGATCACTATGCTTCAAGATCTC
Db	
Oy	241 aggcctgtactacatttgcctgaccctaagattttcatgatacactatgcttcaagatctc
Db	
Oy	301 CATGTCAACAAGATCTCCATGTCAAGATCCAAGTCAGAACAGTCTTCCATCCTCAAGA
Db	
Oy	301 catgtcaacaagatctccatgtcaagatccaagtcagaacaaagcttctccaacctcaaga
Db	
Oy	361 TCTCGATCACAGAGAGAAAATAATGAATGTCAAGGGAAAAAGTAATCTCTGTGATGCTGATT
Db	



Db 1441 ttggttagaataatgtctgacttcaaatgtgatgaaacttgacactatttctaacca 1500

RESULT 4  
AAD26484  
ID AAD26484 standard; DNA; 1500 BP.  
XX  
AC AAD26484;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Mouse alpha (1, 3) galactosyltransferase DNA.  
XX  
KW Mouse; xenotransplantation; Gal-alpha (1,3)Gal; GAL determinant; surgery;  
KW alpha(1,3) galactosyltransferase; alpha (1, 3) GT; drug screening;  
KW gene therapy; ds.  
XX  
OS Mus sp.  
FH  
FH Key Location/Qualifiers  
FT CDS 277..1461  
FT FT /\*tag= a  
FT FT /product= "Mouse alpha (1, 3) GT protein"  
XX  
PN W0200188096-A2.  
XX  
PD 22-NOV-2001.  
XX  
PF 14-MAY-2001; 2001WO-US15765.  
XX  
PR 15-MAY-2000; 2000US-204148P.  
PR 13-JUN-2000; 2000US-0593316.  
XX  
XX (GERO-) GERON CORP.  
XX  
XX Denning C, Clark J;  
XX WPI; 2002-089848/12.  
DR P-PSDB; AAEL6321.  
XX  
PT New ovine tissue devoid of antibody-detectable Gal-alpha(1,3)Gal  
PT determinants, for xenotransplantation, and in the treatment of the  
PT human body by surgery or therapy -  
XX  
PS Claim 16; Page 67-68; 86pp; English.  
XX  
CC The patent discloses immunologically compatible animal tissue, suitable  
CC for xenotransplantation into human patients. The invention also relates  
CC to an ovine tissue devoid of antibody-detectable Gal-alpha (1,3)Gal (GAL)  
CC determinants which are made by alpha(1,3) galactosyltransferase (GT). The  
CC ovine tissue is useful for treatment of human body by surgery or therapy  
CC and in xenotransplantation, by transplanting the ovine tissue into a  
CC mammal having circulating antibody against Gal alpha(1,3)GT determinants.  
CC Polynucleotide constructs of the invention are useful for inactivating  
CC an alpha(1,3) GT gene in an ovine cell. Alpha (1,3) GT sequences are useful  
CC for drug screening and for the production of GAL containing synthetic  
CC oligosaccharides. Sequences of the invention are also useful in gene  
CC therapy. The present sequence is a DNA encoding mouse alpha (1,3) GT  
CC protein.  
XX  
SQ Sequence 1500 BP; 406 A; 349 C; 374 G; 371 T; 0 other;

Query Match 100.0%; Score 1500; DB 24; Length 1500;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTTCCCTTGTACACTCTTCTTGAATGAGAAGTACCGATTCTGCTGAAGACCTCCGCGT 60  
Db 1 ccttccctttagactcttcttggaaatgagaagtaccgattctgctgaagacctgcgct 60  
Qy 61 CTGAGGCTCTGGAGTTGGAAACCTGTACCTTCCCTTCCTCTGCTGAGCCCTGCCCTCCTT 120  
Db 1141 cgggaactgtcggccgcgtacattccaatttcggaagagggtatttttactaccacgcgcgc 1200

Db 61 ctgaggtctcgtggagttggaacctgtaccttcccttctctctgctgagccctgcctcctt 120  
Qy 121 AGGCAGGCCAGAGCTCGACAGAACTCGGTTTGGCTTTGCTGTTTGGAGGGAACACAG 180  
Db 121 aggcagggccagagctccgacagaactcggctgcttctgctgcttctgaggggaaacacag 180  
Qy 181 CTGACGATGAGGCTGACTTTTGAACACTCAAGAGATCTGCTTACCCCCAGTCTCCTCGAATTA 240  
Db 181 ctgacgatgaggtcgtcttgaactcaagagatctgcttaccaccagctcctcctggaattaa 240  
Qy 241 AGGCCTGTACTACATTTGCCTGCACCTAAGATTTTTCATGATCACTATGCTTCAAGATCTC 300  
Db 241 aggcctgtactacatttgcctggaccataagatttcatgatcactatgctcagaatctc 300  
Qy 301 CATGTCAACAAGATCTCCATGTCAAGATCCAAGTCAAGAACTAGAACTCTTCCATCCTCAAGA 360  
Db 301 catgtcaacaagatctccatgtccaagtccaagtcagaacaagctctccatcctccaaga 360  
Qy 361 TCTGGATCACAGAGAAATAATGAATGTCAAGGGAAGAACTAATCTCTGTTGATGCTGATTT 420  
Db 361 tctggatcacagagagaaaaataatgaatgtccaagggaaaaagtaactcctgttgaagtctgatt 420  
Qy 421 GTCTCAACCGTGGTGTCTGCTGTTTGGGAATATGTCAACAGAAATTCAGAGGTTGGTGAG 480  
Db 421 gtctcaacgtggtgtcgtctgttttggaataatgtcaacagaaatccacagaggttggtgag 480  
Qy 481 AACAGATGGCAGAGGACTGCTGTTCCCAAGCTGGTTTAAATAATGGAGCCACAGTTAT 540  
Db 481 aacagatggcagagagactggttggctcccaagctggttcaaaaaatgggacccacagttat 540  
Qy 541 CAAGAAGACAACTAGAAAGGACGAGAGAAAGGGTAGAAATCGAGATCCATTTGAAGAG 600  
Db 541 caagaagacaactagaaaggacgagagaaaaagggtagaaatggagatcgacattgaagag 600  
Qy 601 CCTCAGCTATGGGACTGCTTCAATCCAAAGAACGCCCGGATGTTTTCAGACAGTGACCCCG 660  
Db 601 cctcagctatgggactggttcaatccaaagacccgctggtttttgacagtgaacctcg 660  
Qy 661 TGAAGGCGCGGATTTGTGGGAGGACACTTATGACACAGCTCTGCTGGAAAAAGTACTAC 720  
Db 661 tgaagggcgcgattgtgtgggaaggcacttatgacacagctctgctggaaaagtactac 720  
Qy 721 GCCACAGAAACTACTCTGTGGGGCTGACAGTGTCTGCTGGGAAAGTACATTTGAGCAT 780  
Db 721 gccacagaaactcactgtggggctgacagtgcttctgtgggaaagtacattgagcat 780  
Qy 781 TACTTACAAGACTTTCTGGAGTCTGCTGACATGTACTTTCATGTTGCCCATCGGTCATA 840  
Db 781 tacttgaagacttcttggagctgtctgacatgtacttcatggttggccatcggggtcata 840  
Qy 841 TTTTACGTCTATGACAGACACTCTCCGGATGCCTGCTGACACTGGAACCTCTTACAT 900  
Db 841 ttttaagctatgatagacagacacctccggatgcctgtctgacacctgaacctctacat 900  
Qy 901 TCTTTACAAGCTTTTGAAGTACAGTCTGAGAAGAGGTGGCAGGATATCAGCATGATCGGC 960  
Db 901 tctttacaagtctttagatcaggctcagagagagggtggcaggagatatcagcatgatgcgc 960  
Qy 961 ATGAAGACCAATTGGGGAGGACACATCCTGGCCACATCCAGCAGCAGGTCGACTTCCCTCTTC 1020  
Db 961 atgaagaccattggggagacacacctggccacatccagcaaggctgacttccctcttc 1020  
Qy 1021 TGCATGACCTGGATCAAGTCTTTTCAAGCAAACTTCGGGGTGGAAACTCTGGGCCAGCTG 1080  
Db 1021 tgcattggagctggatcgaagcttcttcaagacaaacttcgggggtggaacctctgggccagctg 1080  
Qy 1081 GTAGCAGCTCCAGGCTCGGTGCTACAGCCAGTCCCGCAGAGTTTTCACCTATGAGAGG 1140  
Db 1081 gtacagactccaggtcgtggttgataaaggccagttcccagagaagttcacctatagagagg 1140  
Qy 1141 CGGGAACCTGTGGCCCGCTACATTCATTCCGAGAGAGGGGATTTTACTACCAGCCGCC 1200  
Db 1141 cgggaactgtcggccgcgtacattccaatttcggaagagggtatttttactaccacgcgcgc 1200

QY 1201 ATTGTTGAGGAACGCCTACTACATTTCTCAACCTCACCAGGAGTGCTTTAAGGGGATC 1260  
|||||  
Db 1201 attttggaggaaacgcctactcacattctcaacctcaccagggagtgtcttaaggggatc 1260  
QY 1261 CTCGAGGACAAGAACATGACATAGAACGCCAGTGGCATGATGAGAGCCACTCAACAAA 1320  
|||||  
Db 1261 ctccaggacaagaacatgacatagaagccagtggtgatgatgagagccacctcaacaaa 1320  
QY 1321 TACTTCCCTTTTCAACAAACCCACTAAATCTATCTCCAGAGTAGTTGCTGGGACTATCAG 1380  
|||||  
Db 1321 tacttctctttcaacaaacccactaaaatctctatctccagagtattgtctgggactatcag 1380  
QY 1381 ATAGCCCTGCCCTCAGATATTTAAAGTGTCAAGGTAGCTTGGCGACAGAAAAGAGATAAT 1440  
|||||  
Db 1381 ataggctgccttcagatatataaagtgtcaaggtagcttggcgacacaaaagataaat 1440  
QY 1441 TTGGTTAGAAATAATGTCTGACTTCAAAATTTGATGTAAGAACTTCACACTATTTCTAACCA 1500  
Db 1441 ttggttagaaataaatgtctgacttcaaatgtgatggaaacttgacactatttctaacca 1500  
RESULT 5  
AAS05315  
ID AAS05315 standard; DNA; 3537 BP.  
XX  
AC AAS05315;  
XX  
DT 07-SEP-2001 (first entry)  
XX  
DE Mouse alpha-1,3 galactosyltransferase genomic sequence #10.  
XX  
KW Mouse; murine; alpha-1,3 galactosyltransferase promoter;  
KW recombinant expression cassette; gene expression; transgenic animal;  
KW livestock; growth hormone; xenotransplantation; transgenic cell; ds.  
XX  
OS Mus musculus.  
XX  
PN WO200130992-A2.  
XX  
PD 03-MAY-2001.  
XX  
PF 20-OCT-2000; 2000WO-US29139.  
XX  
PR 22-OCT-1999; 99US-0161092.  
PR 25-AUG-2000; 2000US-0227951.  
XX  
PA (UYPI-) UNIV PITTSBURGH.  
XX  
PI Kolke C;  
XX  
DR WPI; 2001-300505/31.  
XX  
PT New recombinant expression cassette useful in xenotransplantation -  
PS Claim 16; Page 77-80; 138pp; English.  
XX  
CC The present sequence for mouse alpha-1,3 galactosyltransferase (GT)  
CC genomic sequence #10 is 1 of 44 alpha-1,3 galactosyltransferase  
CC gene sequences (AAS05290-AAS05333) from different animals. The  
CC invention provides a novel recombinant expression cassette comprising  
CC an alpha-1,3 GT promoter operably linked to a polynucleotide for  
CC expression. The cassettes of the invention can be used to express  
CC foreign genes or to disrupt the native alpha-1,3 GT genomic sequence.  
CC The invention provides transgenic animals and methods for their  
CC production. Transgenic mice in which a reporter gene is operably linked  
CC to the alpha-1,3 GT promoter can be used for assessing promoter activity  
CC and specificity. Transgenic livestock having the novel expression  
CC cassette in which a growth hormone is expressed under the control of the  
CC alpha-1,3 GT promoter can be matured or grown better than commonly  
CC employed strains. Tissue obtained from transgenic animals can be  
CC implanted into a host providing a method of xenotransplantation from a  
CC transgenic animal. The invention also provides a transgenic organ

CC consisting essentially of transgenic cells engineered, for use in  
CC xenotransplantation. The methods of the invention facilitate  
CC exhibiting differential expression of the gal-alpha-gal epitope. The  
CC invention also describes primers (AAS05334-AAS05385) used to isolate the  
CC alpha-1,3 GT sequences.  
XX  
SQ Sequence 3537 BP; 978 A; 777 C; 843 G; 939 T; 0 other;  
  
Query Match 95.7%; Score 1434.8; DB 22; Length 3537;  
Best Local Similarity 97.5%; Pred. No. 0;  
Matches 1493; Conservative 0; Mismatches 2; Indels 37; Gaps 2;  
  
QY 1 CCTTCCCTTGTAGACTCTTCTTGGAAATGAGAAGTACCGATTCTGCTGAAGACCTCGCGCT 60  
|||||  
Db 120 ccttcccttgtagactcttcttggaaatgagaagtaacgattctgtgaagacctcgcgt 179  
QY 61 CTCAGGCTCTGGGAGTTGGAACCTGTACCTTCTTCCCTGCTGAGCCCTGCTCTT 120  
|||||  
Db 180 ctcaaggctctgggagtggaaacctgtaccttcttctctctgtgagccctgctctt 239  
QY 121 AGCAGGCCAGAGCTCGACAGAA-CTCGGTTGCTTTTGTGTTTGGAGGAAACACA 179  
|||||  
Db 240 aggcaggccagagctcgacagaagctcggttgctgttcttcttggaggagaaca 299  
QY 180 GCTCAGCATGAGCTGACTTTTGAATCAAGAGATCTGCTTACCCAGTCTCCTCGAATTA 239  
|||||  
Db 300 gctcagcatgagctgacttggacttgaactcaagagatctgttaccctcagctccttggaa 359  
QY 240 AAGCCCTGTACTACATTTGCTGCGACCTTAAGATTTTCATGATCACTATGCTTCAAGATCT 299  
|||||  
Db 360 aaggccgtgactacattgctggtgacctaaagtattcatgatcactatgcttcaagatct 419  
QY 300 CCATGTCAACAAGATCTCCATGTCAAGATCCAAGTCAAGTCAAGAAACAAGTCTTCCATCTCAAG 359  
|||||  
Db 420 ccatgtcaacaagatctccatgtcaagatccaaagtcgaagaagaagtttccatcctcgaag 479  
QY 360 ATCTGGATCACAGAGAAATAATGAATGTCAAGGAAAGAACTAATCTCTGTGTGATGCTGAT 419  
|||||  
Db 480 atctggatcacagagagaaaaaaataatgaatgtcaagggaagtaactcctgttgaatgctgat 539  
QY 420 TGTCTCAACCGTGGTGTCTGCTGTTTGGGAATATGTCAACAG----- 461  
|||||  
Db 540 tgtctcaacctggtgtcgtgttcttggaaatatgtcaacagccccagagcgtcttctt 599  
QY 462 -----AATTCCAGAGGTTGGTGAGAACAGATGGCAGAGAGGACTGGTG 503  
|||||  
Db 600 gtggatatatacacacaaaaattccagaggttgggtgagacagatggcagaagagactggtg 659  
QY 504 GTTCCCAAGCTGGTTTAAATGGGACCCACAGTTTATCAAGAAGACAACGTTAGAGGACG 563  
|||||  
Db 660 gtcccaagctggtttaaaaatgggacccacagttaataagaagaacaacgtagaaggacg 719  
QY 564 GAGAAAGGTTAGAAATGGAGATCGCATTTGAAGCCCTCAGCTATGGGACTGGTTCAA 623  
|||||  
Db 720 gagaaagaaggtagaataaggagatcgcattgaagagccctcagctatgggactggttcaa 779  
QY 624 TCCAAAGAACCGCCCGGATGTTTTCACAGTGCACCCCTGGAAAGCGCCGATTTGTGGGA 683  
|||||  
Db 780 tccaaagaacccgcccggatgttttgacagtgacccctgggaaggcgcgattgttctggga 839  
QY 684 AGGCATTTATGACACAGCTCTGTGGAAAGTACTAGCCACACAGAAACTCATCTGTGGG 743  
|||||  
Db 840 aggcacttatgacacagctctgtggaagaagtactacgcccacacagaaactcactgtgg 899  
QY 744 GCTCAGAGTGTTCCTGTGGGAAAGTACATTTGACATTACTTTAGAAGACTTTCTGGAGTC 803  
|||||  
Db 900 gctgacagtgcttctgctgtggaaagtacatttgagcatcacttagaagactcttcggagtc 959  
QY 804 TGTGACATGTACTTTCATGTTGGCCATCGGGTTCATATTTTACGTTCATGTATGACGACAC 863  
|||||  
Db 960 tgcctgacatgtacttcatcgttggccatcgggttcataattttacgtcatgatgatgacac 1019

QY 864 CTCCCGATGCTCTGTCACCTGACCTGAACTTACATCTCTTACAGTCTTTGAGATCAG 923  
DB 1020 CTCGCGATGCTCTGTCACCTGACCTGAACTTACATCTCTTACAGTCTTTGAGATCAG 1079  
QY 924 GTCTGAGAAAGTGGCAGGATATCAGCATGATGCCCATGAAGACCATTTGGGAGACAT 983  
DB 1080 GTCGGAAGAGTGGCAGGATATCAGCATGATGCCCATGAAGACCATTTGGGAGACAT 1139  
QY 984 CTGGGCCACATCCAGCAGGATGCTCTCTCTCTGATGAGCTGATCAAGCTCT 1043  
DB 1140 CTTGGCCACATCCAGCAGGATGCTCTCTCTCTGATGAGCTGATCAAGCTCT 1199  
QY 1044 TCAAGCAACTTCGGGGTGAAGTCTGGGCCAGTGTGGTACACACTCCAGGCCCTGGTG 1103  
DB 1200 TCAAGCAACTTCGGGGTGAAGTCTGGGCCAGTGTGGTACACACTCCAGGCCCTGGTG 1259  
QY 1104 GTACAGGCCAGTCCAGCAAGTTCACCTATGAGAGCGGGAACTCTCGGCCCGTACAT 1163  
DB 1260 GTACAGGCCAGTCCAGCAAGTTCACCTATGAGAGCGGGAACTCTCGGCCCGTACAT 1319  
QY 1164 TCCATTTCGGAGAGGGGATTTTACTTACCACCGCGGCATTTTGGAGGAACGCTACTCA 1223  
DB 1320 TCCATTTCGGAGAGGGGATTTTACTTACCACCGCGGCATTTTGGAGGAACGCTACTCA 1379  
QY 1224 CATTTCAACCTCACAGGGAGTGTCTTAAAGGGATCTCTCCAGGACAAACATGACAT 1283  
DB 1380 CATTTCAACCTCACAGGGAGTGTCTTAAAGGGATCTCTCCAGGACAAACATGACAT 1439  
QY 1284 AGAAGCCAGTGCATGATGAGAGCCACCTCAACAATACTCTCTTTCAACAACCCAC 1343  
DB 1440 AGAAGCCAGTGCATGATGAGAGCCACCTCAACAATACTCTCTTTCAACAACCCAC 1499  
QY 1344 TAAATCTATCTCCAGAGTATTGCTGGGACTATCAGATAGGCTGCTCTCAGATATTAA 1403  
DB 1500 TAAATCTATCTCCAGAGTATTGCTGGGACTATCAGATAGGCTGCTCTCAGATATTAA 1559  
QY 1404 AAGTGTCAAGGTAGCTTGGCAGACAAAGAGTATATTTGGTTAGAAATAATCTCTGACT 1463  
DB 1560 AAGTGTCAAGGTAGCTTGGCAGACAAAGAGTATATTTGGTTAGAAATAATCTCTGACT 1619  
QY 1464 TCAAAATGTGATGGAACCTTGACACTATTCT 1495  
DB 1620 TCAAAATGTGATGGAACCTTGACACTATTCT 1651

## RESULT 6

AAT04523  
ID AAT04523 standard; cDNA; 1116 BP.

AC AAT04523;

XX 01-MAY-1996 (first entry)

DE Mouse alpha-1,3-galactosyltransferase coding region.

XX Mouse; alpha-1,3-galactosyltransferase; Immune response; glycoprotein;  
KW alpha-galactosyl epitope; cell membrane; virus; phagocytosis; tumour;  
KW antigen processing; leukemia; lymphoma; myeloma; melanoma; carcinoma;  
XX sarcoma; vaccine; opsonisation; glycoprotein; antibody; anti-Gal; ss.

OS Mus musculus.

XX WO9524924-A1.

XX 21-SEP-1995.

XX 13-MAR-1995; 95WO-US03156.

XX 15-MAR-1994; 94US-0213200.

XX (UYHA-) UNIV HAHNEMANN &amp; MEDICAL COLLEGE PENNSYL.

XX

PI Gallili U, Replik PM;

XX WPI; 1995-336816/43.

XX Association of an alpha-galactosyl epitope with a tumour or viral  
PT antigen - is administered to anti-Gal synthesising animals to induce  
PT an immune response

PS Disclosure; Fig 9; 85pp; English.

XX The nucleotide sequence of the mouse alpha-1,3-galactosyltransferase  
CC coding sequence. The enzyme can be used in methods of enhancing an  
CC immune response by associating the alpha-galactosyl epitope with a cell  
CC membrane or viral glycoprotein. The alpha-galactosyl epitope enhances  
CC phagocytosis and subsequent processing of the antigen. The method is  
CC useful in the treatment of tumours e.g. leukemia, lymphoma, myeloma,  
CC melanoma, carcinoma and sarcoma, or for the generation of viral vaccines  
CC by opsonising a viral glycoprotein. The alpha-galactosyl epitope  
CC enhances recognition of the antigen in an animal that synthesises the  
CC naturally occurring antibody - anti-Gal.  
CC See AAT04522 for the marmoset alpha-1,3-galactosyltransferase sequence.

SQ Sequence 1116 BP; 309 A; 250 C; 293 G; 264 T; 0 other;

Query Match 68.7%; Score 1030.8; DB 16; Length 1116;  
Best Local Similarity 96.6%; Pred. No. 3e-310;  
Matches 1078; Conservative 0; Mismatches 2; Indels 36; Gaps 1;

QY 382 ATGAATGTCAGGGAAGTAATCTCTGTGATGCTCTGATCTCTCAACCGTGTGCTG 441

DB 1 atgaatgcaagggaagtgatcctgttgatgctgattgtctcaacccgtggtgtcgtg 60

QY 442 TTTTGGGAATATCTCAACAG-----AATT 465

DB 61 ttttgggaatatctcaacagccagagcgtcttcttgggtatatacaaaaaatt 120

QY 466 CCAGAGGTTGGTGAGAACAGATGGCAGAGGACTGTGGTTCCTCAAGCTGGTTAAAT 525

DB 121 ccagaggttggtgagaaagatggcagagagctggtgtcccaagctggtttaaaaat 180

QY 526 GGGACCCACAGTTATCANGAAGACACGTAGAGGACGGAGAGAGGAGGATAGATGGA 585

DB 181 gggacccacagttatcaagaagacaacgtagaaggcggagagaagaggtagaatgga 240

QY 586 GATCGCATTTCAAGAGCTCTAGCTATGGAGCTTCAATCCAAAGAACCCCGGATGTT 645

DB 241 gatcgcattgaaagacctcaagtattggagctggttcaatccaaagaaccccgagatgtt 300

QY 646 TTGACAGTGACCCCGTGGAAGCGCCGATTTGTGTGGGAAGGCACATTATGACAGCTTG 705

DB 301 ttgacagtgaccccggtggaagcgccgattgtgtggaagggcacttatgacacagctctg 360

QY 706 CTGGAAGTACTACGCCACACAGAACTCTCTGGGGCTGACAGTGTGCTGTGCGGA 765

DB 361 ctggaagttactacgccacacagaaactcactgtgggctgacagtgctgtgtgtgga 420

QY 766 AACTACATTGAGCATTTACTTAGAAGACTTCTCGAGTCTCTGACATGTACTTTCATGTT 825

DB 421 aagtacatgagcatctacttagaagacttctctgagctgtgagctgacatgactcatggtt 480

QY 826 GGCCATCGGTCATATTTTACGTATGATAGACGACACCTCCCGGATGCGTCTGTCAC 885

DB 481 ggcctatcggtcatatttttaagtcaltgatagacacacctcccgatgctgtcgtgac 540

QY 886 CTGAACCTCTTACATCTCTTACAGTCTTTGAGATCAGTCTGAGAGAGGTCGCCAGAT 945

DB 541 ctgaacctctacattccttcaagctctttgagatcaggtctgagaagaggtgagagat 600

QY 946 ATCAGCATGATGCGCATGAAGACCATTTGGGAGCAGACATCTGCCCCACATCCAGCAG 1005

DB 601 atcagcatgatgctgatgaagacatttggggagcacatctctggccacatccacgacgag 660

QY 1006 GTCGACTTCTCTTCATGAGCTGGATCAAGCAACTCTTCGAGCAAACTTCGGGGTGGAA 1065  
DB 661 gtgacttctcttctgcatgagctggatcaagctcttcaagacaactctcggggtggaa 720  
QY 1066 ACTCTGGCCAGCTGGTAGCACAGCTCCAGGCTGGTGTACAGGCGAGTCCCGAGAA 1125  
DB 721 actctggccagctggtagcacagctccaggtcggtggtagaaggccagtcgccgaac 780  
QY 1126 TTCACCTATGAGAGCGGGGAACCTCGCGCGGTACATTCATTCGGAGAGGGGATTTT 1185  
DB 781 ttcacclatgagagcggaactctcgccgcgtacattccattcgcgaggggagatttt 840  
QY 1186 TACTACCACGGCCATTTTGGAGAACGCTACTCACATTCCTCAACCTCACAGGAG 1245  
DB 841 tactaccacgcgccattttggaggacgctactcacattctcaacctcaccaggag 900  
QY 1246 TGCTTTAAGGGATCTCCAGGACAAGAAACATGACATAGAAGCCAGTGGCATGATGAG 1305  
DB 901 tgctttaagg99ga tccctccg99gaagaacatgacatagaagcccg9tgcac tga tga 960  
QY 1306 AGCCACCTCAACAAATACTTCTCTTTCAACAAACCCACTAAAATCCTATCTCCAGAGTAT 1365  
DB 961 agccacctcaacaatactctcttttcaacaacccactaaaaatcctatctccagagtat 1020  
QY 1366 TGCTGGACTATCAGATAGGCTGCTCCTCAGATATTAAAAGTGTCAAGGTAGCTTGGCAG 1425  
DB 1021 tgctggactatcagatag9ctgcctcagatattaaaagtgtcaaggtagcttggcag 1080  
QY 1426 ACAAGAGTATATTTGGTTAGAAATAATCTCTGA 1461  
DB 1081 acaaaagatataatttgg9ttagaaa caatg9ctga 1116

RESULT 7

AAQ74712  
ID AAQ74712 standard; cDNA to mRNA; 1423 BP.  
XX  
AC AAQ74712;  
XX  
DT 26-JUN-1995 (first entry)  
XX  
DE Galactosyl transferase clone.  
XX  
KW Gal-alpha (1,3) galactosyl transferase; xenograft; transplant;  
KW rejection; ss.  
XX  
OS Sus scrofa domestica.  
XX  
FH Key Location/Qualifiers  
FT CDS 50..1129  
FT /\*tag= a  
XX  
XX WO9421799-A.  
XX  
XX 29-SEP-1994.  
XX  
XX 15-MAR-1994; 94WO-AU00126.  
XX  
XX 16-MAR-1993; 93AU-0007854.  
XX  
XX (AUST-) AUSTIN RES INST.  
XX  
XX Mckenzie IFC, Sandrin MS;  
XX  
XX WPI; 1994-317019/39.  
XX  
XX P-PSDB; AAR62508.  
XX  
XX DNA sequences encoding Gal-alpha (1,3)galactosyl transferase -  
XX and clones contg. such sequences are used in xenograft therapies  
XX  
XX Claim 3; Page 35; 50pp; English.  
XX  
XX The sequence is that of the porcine Gal-alpha (1,3) galactosyl

CC transferase gene which produces a Gal epitope on the surface of  
CC porcine cells. This epitope is recognised by antibodies which are  
CC responsible for hyperacute rejection of xenotransplanted pig cells,  
CC tissues and organs  
CC See also AAQ74711-4.  
XX  
SQ Sequence 1423 BP; 413 A; 305 C; 348 G; 357 T; 0 other;

Query Match 48.4%; Score 726.6; DB 15; Length 1423;  
Best Local Similarity 78.8%; Pred. No. 2.1e-215; Indels 6; Gaps 2;  
Matches 893; Conservative 0; Mismatches 234;  
QY 354 CTCAAGATCTGGATCACAGGAGAAATAATGAATGTCAAGGAAAGTAATCTCTGTGAT 413  
DB 22 cccagctctcgccatcaggagaaaaatgaatgtcaaggagagtggtctcgtcgaat 81  
QY 414 GCTGATTTGCTCAACCGTGGTTGCTGTTGGGAATATGTCAACAGATTCACAGAGT 473  
DB 82 gctgcttctcaactgaatggttcttgggaatacatcaacagaaaccagagagt 141  
QY 474 TGGTGAGAACAGATGGCAGAGGACTGGTGTCCCAAGCTGGTTTAAATAATGGAGCCA 533  
DB 142 tgg ---cagcagtgctcagaggggctggtgttcccgagctggttcaacaatggagactca 198  
QY 534 CAGTTATCAAGAAGACAACCTAGAGGACGAGAGAAAAGG---GTAGAAATGGAGATCG 590  
DB 199 cagttaccagagagagagacgcctataggcaacgaaaggaaacaaagaaagagacaa 258  
QY 591 CATTGAAGAGCCTCAGTATGGGACTGGTTCAATCCAAAGAACCCGCGGATGTTTGAC 650  
DB 259 cagaggagagcttcgcgtagtggaactggtttaaactcgagaaacgcccagaggtcgtgac 318  
QY 651 AGTACCCCGTGGAAAGCGCGGATTTGTGGGAGGCACTTATGACACAGCTCTGCTGGA 710  
DB 319 cataaccagatggaaaggtccagtggtatgggaaggcacttaacaacagagcgctctaga 378  
QY 711 AAAGTACTACCCACACAGAAACTCACTGTGGGCTGACAGTGTGTTGCTGTGGAAAGTA 770  
DB 379 taattattatgcaaacagaaaaattaccgtgggcttgagcgggttttctgctgcggaagata 438  
QY 771 CATTGACATTTACTTAGAGACTTTCTGGAGTCTGCTGACATGTACTTCATGGTTGGCCA 830  
DB 439 cattgagacttaactggagagttcttaatactctgcaaatatactacttcatggttggcca 498  
QY 831 TCGGTCATATTTACGTCATGATGACAGCACCTCCCGGATGCTGCTGCACTGAA 890  
DB 499 caaagtcabctctttacatcatatggatgatatctccaggatgcctttgatagagctggg 558  
QY 891 CCCTCTACATTCCTTACAAGTCTTTTGAGATCAGGTCTGAGAAAGAGGTGGCAGGATACAG 950  
DB 559 tcctctgcttctcttaaagtggttgagatcaagtcgagaaagaggtggcaagacatcag 618  
QY 951 CATGATGGCATGAAGACCATTTGGGGAGCACATCTTGGCCACATCCAGCAGAGGTGCA 1010  
DB 619 catgatcgcatgaagaccatcgggagcacatctctggccacatccagcagcagctggg 678  
QY 1011 CTTCCTCTTCTGCAATGGACGTGATCAAGTCTTTCAAGACAACCTTCGGGGTGGAACTCT 1070  
DB 679 ctctctctctgcaatgacgtggaataggtcttccaaaacaactttggggtggagaccct 738  
QY 1071 GGGCCAGCTGGTAGCAGCTCCAGGCTGGTGTGTACAGGCCAGTCCCGAGAGATTAC 1130  
DB 739 gggccagtcggtggctcagctacagggcctggtgtacaaagcacatcctcagcaggttcac 798  
QY 1131 CTATGAGAGGGGGAACACTTCGGCCCGCTACATTCATTCGAGAGGGGATTTTACTA 1190  
DB 799 ctacgagagcggaaggagtcgcagctacatctccgttttggccaggggagattttatta 858  
QY 1191 CCAGCGCGCCATTTTGGAGGAAGCCCTACTTCATATCTCAACCTCACCAGGGAGTGT 1250  
DB 859 ccagcagccatttttggggaacacccactcaggttcttaaacatcactcagcaggtgctt 918

	Matches	893;	Conservative	0;	Mismatches	234;	Indels	6;	Gaps	2;
Qy	354	CTCAAGATCTCGATC	CACAGGAGAAATAAT	CAATGTCTCAAGGCAAAAGTAAT	CCCTGTTGAT	413				
Db	22	cccaagcttctgcgatc	caggagaaaaataat	gaatgtcctcaaggaagag	tggtctctgtcgaat	81				
Qy	414	GCTGATTGTCTCAACC	GGTGTGTCGTGTTT	TGGGAATATCTCAACAGAAAT	TTCACAGAGT	473				
Db	82	gctgcttgtctcaactg	taatgggttggtgtgt	tttgggaaatcatc	caacagaccacagagag	141				
Qy	474	TGGTGTAGACACAGAT	TGGCAGAGGACTGGT	GTGTCCCAAGCTGGT	TTAAAAATGGGACCCA	533				
Db	142	tgg---cagcagtgctc	agagggctgggtgt	ttccagagctgggt	tttaacaatggagactca	198				
Qy	534	CAGTTATCAACAGACA	CAACGTAGAGAGGAC	GAGAGAGAAAGG--	-GTACAAATGGAGATCG	590				
Db	199	cagttaccacgaagaa	gaagacgctatagg	caacgaagaaagaa	gaagaaagaaagaa	258				
Qy	591	CATTGAAGAGCCTCAG	CTATGGGACTGGT	GTCAATCCAAAGAAC	CGCCCGGATGTTT	650				
Db	259	cagaggagagcttcgc	gtagtggaactgggt	tttaactctgagaa	agccacagaggtcgctgac	318				
Qy	651	AGTGACCCGTGGAAG	CGCCGATGTGTGG	AAGGCACATTAT	GACACAGACTCTCTCTGGA	710				
Db	319	cataacaga	tggaaggtccag	tggtcatgggaagggca	ttcacacagagcgctctaga	378				
Qy	711	AAAGTACTACGCCAC	ACAGAAACTCAC	TGTGGGCTGACAGT	GTTGTGTGGGAAGTA	770				
Db	379	taattattatgcca	aaacagaaaaat	taccgtgggtctgag	gttttctgtgtcggaagata	438				
Qy	771	CATTGAGCAITTTAC	TTAGAAAGACTT	TCTGGAGTCTCTG	ACATGTACTTCAATGGTGTGGCCA	830				
Db	439	catlgagcat	tacttgaggaggt	tcttaatatctg	caaatatactacttca	498				
Qy	831	TCGGGTATATTTTAC	GTTCATGATAGAC	GACACCTCCGCGAT	GCCTGTCTGTCCACCTCAA	890				
Db	499	caaaatcatctttaca	tcatctgttgga	tgaatctccaggtatg	cttttga	558				
Qy	891	CCCTCTACATCTCTT	ACAACTCTTTGAG	ATCAGGTCTCAGA	AGAGTGGCAGGATATCAG	950				
Db	559	tcctctgcgttcc	ttttaaagtgttt	tgagatacaagt	ccgagaagaggtggc	618				
Qy	951	CATGATCGCATGAAG	ACCATTTGGGGAG	CACATCCTTGCC	CCACATCCAGCACGAGTCCA	1010				
Db	619	catgatgcgatgaag	acca	tcggggagcacat	ccttggtccacatccag	678				
Qy	1011	CTTGCTCTTCTGCAT	GTGAGCTGGATCA	AGTCTTTCAACACA	CACTTCGGGGTGGAACTCT	1070				
Db	679	ctctctctctgcat	tgagtgagtcaggt	ctctccaaacact	tttgggttggaacct	738				
Qy	1071	GGGCGAGTGTAGCAC	AGCTCCAGGCTGT	GTGTACAGGCC	AGTCTCCGAGAGCTTCAC	1130				
Db	739	gggcagtcggtcgct	agctacagggctcg	tggaacaaagaca	ctctcagcagctcac	798				
Qy	1131	CTATGTAGAGGGGA	AACTCTCGCGCGCT	ACATCTCCAT	TTCGGAGAGGGGATTTTTACTA	1190				
Db	799	ctacagaagcggaag	agagtcgcgac	ctacatccgt	tttgccaggggatttttata	858				
Qy	1191	CCACGGGCCCATTTT	TGGAGGAACGCCT	TACTTCACAT	TCTCAACCTCACCGGAGTGTCT	1250				
Db	859	ccaagcaccat	ttttgggggaa	caccacactcag	gttttttataaca	918				
Qy	1251	TAAAGGATCTCCAG	GACAAGAACAT	GACATACA	AGCCAGTGGCATGATGAGAGCCA	1310				
Db	919	caaaggaa	tcctccagga	caggaaaa	tgacataagaa	978				
Qy	1311	CCTCAACAATATCT	CTCTTTTACAAAC	CCCACTAAAA	TCTATCTCCAGAGTATTTGCTG	1370				
Db	979	tctaaacaagat	ttctcttca	aaaccccaataa	ttcttataccccaagaatactgctg	1038				
Qy	1371	GGACTATCAGATAG	CGCTTTCAGAT	ATTAAAA	GTCTCAGGTAGCTTGGCAGACAAA	1430				
Db	1039	qgaattatcata	taggca	tgctgtgga	tatttaggattctcaaga	1098				

RESULT	8
AAT12242	
ID	AAT12242 standard; cDNA to mRNA; 1423 BP.
XX	
XX	
AC	AAT12242;
XX	
DT	08-APR-1996 (first entry)
XX	
DE	Pig alpha(1,3)-galactosyltransferase, full coding sequence.

Qy	711	AAAGTACTACGCCACACAGAAACTCACTGTGGSGCTGCACAGTGTTCCTCGGAAAGTA	770
Db	379	taatttatatgcccaacagaaaaattaccgtggagcttgagcgttttttgctatcggaagata	438

771 CATTCAGCGCATATACATTAGAGAGACATTATCTCGAGCTCTGTCGACACATGTAATTCAATCGTTGGCCA 830

Db	439	catcgagcatcactgggaggttctcaatatctgcaatatatacttcatggttgccca	498
QV	831	TCGGGTTCATATTTTACCTTCATCATAGACGACACCTCCCGGATGCCTCGTCGCACCTGA	890

Db 499 caaagtcatctttttacatcatggtgaatgaatattctccaggatgccctttgaatagaacctga 558

891 CCTCTACATTCCTTACAAGTCTTTGAGATCAGGTCCTGAGAAGAGAGGTCGCAGGATATCAG 950

[illegible][illegible]

1. **Introduction**  
 2. **Background**  
 3. **Methodology**  
 4. **Results**  
 5. **Discussion**  
 6. **Conclusion**  
 7. **References**  
 8. **Appendix**  
 9. **Figure 1**  
 10. **Figure 2**  
 11. **Figure 3**  
 12. **Figure 4**  
 13. **Figure 5**  
 14. **Figure 6**  
 15. **Figure 7**  
 16. **Figure 8**  
 17. **Figure 9**  
 18. **Figure 10**  
 19. **Figure 11**  
 20. **Figure 12**  
 21. **Figure 13**  
 22. **Figure 14**  
 23. **Figure 15**  
 24. **Figure 16**  
 25. **Figure 17**  
 26. **Figure 18**  
 27. **Figure 19**  
 28. **Figure 20**  
 29. **Figure 21**  
 30. **Figure 22**  
 31. **Figure 23**  
 32. **Figure 24**  
 33. **Figure 25**  
 34. **Figure 26**  
 35. **Figure 27**  
 36. **Figure 28**  
 37. **Figure 29**  
 38. **Figure 30**  
 39. **Figure 31**  
 40. **Figure 32**  
 41. **Figure 33**  
 42. **Figure 34**  
 43. **Figure 35**  
 44. **Figure 36**  
 45. **Figure 37**  
 46. **Figure 38**  
 47. **Figure 39**  
 48. **Figure 40**  
 49. **Figure 41**  
 50. **Figure 42**  
 51. **Figure 43**  
 52. **Figure 44**  
 53. **Figure 45**  
 54. **Figure 46**  
 55. **Figure 47**  
 56. **Figure 48**  
 57. **Figure 49**  
 58. **Figure 50**  
 59. **Figure 51**  
 60. **Figure 52**  
 61. **Figure 53**  
 62. **Figure 54**  
 63. **Figure 55**  
 64. **Figure 56**  
 65. **Figure 57**  
 66. **Figure 58**  
 67. **Figure 59**  
 68. **Figure 60**  
 69. **Figure 61**  
 70. **Figure 62**  
 71. **Figure 63**  
 72. **Figure 64**  
 73. **Figure 65**  
 74. **Figure 66**  
 75. **Figure 67**  
 76. **Figure 68**  
 77. **Figure 69**  
 78. **Figure 70**  
 79. **Figure 71**  
 80. **Figure 72**  
 81. **Figure 73**  
 82. **Figure 74**  
 83. **Figure 75**  
 84. **Figure 76**  
 85. **Figure 77**  
 86. **Figure 78**  
 87. **Figure 79**  
 88. **Figure 80**  
 89. **Figure 81**  
 90. **Figure 82**  
 91. **Figure 83**  
 92. **Figure 84**  
 93. **Figure 85**  
 94. **Figure 86**  
 95. **Figure 87**  
 96. **Figure 88**  
 97. **Figure 89**  
 98. **Figure 90**  
 99. **Figure 91**  
 100. **Figure 92**  
 101. **Figure 93**  
 102. **Figure 94**  
 103. **Figure 95**  
 104. **Figure 96**  
 105. **Figure 97**  
 106. **Figure 98**  
 107. **Figure 99**  
 108. **Figure 100**  
 109. **Figure 101**  
 110. **Figure 102**  
 111. **Figure 103**  
 112. **Figure 104**  
 113. **Figure 105**  
 114. **Figure 106**  
 115. **Figure 107**  
 116. **Figure 108**  
 117. **Figure 109**  
 118. **Figure 110**  
 119. **Figure 111**  
 120. **Figure 112**  
 121. **Figure 113**  
 122. **Figure 114**  
 123. **Figure 115**  
 124. **Figure 116**  
 125. **Figure 117**  
 126. **Figure 118**  
 127. **Figure 119**  
 128. **Figure 120**  
 129. **Figure 121**  
 130. **Figure 122**  
 131. **Figure 123**  
 132. **Figure 124**  
 133. **Figure 125**  
 134. **Figure 126**  
 135. **Figure 127**  
 136. **Figure 128**  
 137. **Figure 129**  
 138. **Figure 130**  
 139. **Figure 131**  
 140. **Figure 132**  
 141. **Figure 133**  
 142. **Figure 134**  
 143. **Figure 135**  
 144. **Figure 136**  
 145. **Figure 137**  
 146. **Figure 138**  
 147. **Figure 139**  
 148. **Figure 140**  
 149. **Figure 141**  
 150. **Figure 142**  
 151. **Figure 143**  
 152. **Figure 144**  
 153. **Figure 145**  
 154. **Figure 146**  
 155. **Figure 147**  
 156. **Figure 148**  
 157. **Figure 149**  
 158. **Figure 150**  
 159. **Figure 151**  
 160. **Figure 152**  
 161. **Figure 153**  
 162. **Figure 154**  
 163. **Figure 155**  
 164. **Figure 156**  
 165. **Figure 157**  
 166. **Figure 158**  
 167. **Figure 159**  
 168. **Figure 160**  
 169. **Figure 161**  
 170. **Figure 162**  
 171. **Figure 163**  
 172. **Figure 164**  
 173. **Figure 165**  
 174. **Figure 166**  
 175. **Figure 167**  
 176. **Figure 168**  
 177. **Figure 169**  
 178. **Figure 170**  
 179. **Figure 171**  
 180. **Figure 172**  
 181. **Figure 173**  
 182. **Figure 174**  
 183. **Figure 175**  
 184. **Figure 176**  
 185. **Figure 177**  
 186. **Figure 178**  
 187. **Figure 179**  
 188. **Figure 180**  
 189. **Figure 181**  
 190. **Figure 182**  
 191. **Figure 183**  
 192. **Figure 184**  
 193. **Figure 185**  
 194. **Figure 186**  
 195. **Figure 187**  
 196. **Figure 188**  
 197. **Figure 189**  
 198. **Figure 190**  
 199. **Figure 191**  
 200. **Figure 192**  
 201. **Figure 193**  
 202. **Figure 194**  
 203. **Figure 195**  
 204. **Figure 196**  
 205. **Figure 197**  
 206. **Figure 198**  
 207. **Figure 199**  
 208. **Figure 200**  
 209. **Figure 201**  
 210. **Figure 202**  
 211. **Figure 203**  
 212. **Figure 204**  
 213. **Figure 205**  
 214. **Figure 206**  
 215. **Figure 207**  
 216. **Figure 208**  
 217. **Figure 209**

[illegible]

QY 10/I GGGCCAGCTGGTAGCACAGCTCCAGGCCCTGGTGGTACAGGCCAGTCCCGAAGATCAG 1130

Db 739 gggccagtcggctggctcagctacaggccctgggtggtacaaaggcacatcctgacgagttcac 798

Qy 1131 CTATGAGAGCGGGAACGTCTCGGCCCGGTACATTCCATTCCGAGAGGGGATTTTTACTA 1190

Db 799 ctacgagggcgaaggagtcgcgagcctacattccgtttggccaggggattttatta 858

QY 1191 CCACGGGCCATTTTTGGAGGAACGCCCTACTCACATTCTCAACCTCACCGGGAGTGCTT 1250

Db 859 ccacgcagccatttttgggggaacacccactcaggattctaaacatcactcaggaatgctt 918

09 1251 TAAGGGGATCCTCCAGGACAAGAAAACATGACATAGAGGCCAGTGGCATGATGAGAGGCCA 1310

[illegible]

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

-----

QY 1571 00AC1A1CAGRIAG00C1GCCI1CAGATA11A1AAAG101C1A1AG01AGC1100CAGACAAA 1430

Db 1039 ggattatcatataggaatgctctgtggaatctaggaatctcagaatagccttggaagaaaa 1098



DE Porcine alpha-1,3-galactosyl transferase isoform 2 cDNA.  
XX  
KW Isoform: porcine; enzyme: alpha-1,3-galactosyl transferase; galactose;  
KW sugar; N-acetyllactosamine; glycoprotein; glycolipid; antibody; pig;  
KW graft tissue rejection; organ transplantation; xenotransplant; ss.  
XX  
OS Sus scrofa.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1090  
FT /tag= a  
FT /product= "alpha-1,3-galactosyl transferase isoform 2"  
FT /note= "contains no start or stop codon at the 5' or 3'  
FT ends of the sequence"  
XX  
PN FR2751346-A1.  
XX  
PD 23-JAN-1998.  
XX  
PF 19-JUL-1996; 96FR-0009077.  
XX  
PR 19-JUL-1996; 96FR-0009077.  
XX  
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
XX  
PI Pourcel C, Soullillon JP, Vanhove B;  
XX  
DR WPI; 1998-112876/11.  
XX  
DR P-PSDB; AAW49687.  
XX  
PT Transgenic non-human donors of organs for human recipients -  
PT containing DNA encoding antibodies that inhibit graft rejection  
XX  
PS Claim 4; Page 36-37; 71pp; French.  
XX  
CC This sequence represents the cDNA encoding isoform 2 of the porcine  
CC enzyme alpha-1,3-galactosyl transferase (alpha-1,3-GT). The enzyme  
CC catalyses the attachment of a galactose sugar molecule on the  
CC N-acetyllactosamine moiety found on surface glycoproteins and  
CC glycolipids. These sugar molecules are partly responsible for raising  
CC anti-graft antibodies, which lead to graft tissue rejection. The  
CC invention relates to a method of inhibiting the graft rejection mechanism  
CC by introducing the sequence encoding an antibody targeted to alpha-1,3-GT  
CC into the cells of animal, especially a pig, from whom organs may be used  
CC for xenotransplants. Neutralisation of the alpha-1,3-GT leads to tissues  
CC or organs lacking the galactose on the glycoproteins and glycolipids,  
CC thus preventing induction of the rejection response.  
SQ Sequence 1092 BP: 325 A; 228 C; 279 G; 260 T; 0 other;

Query Match 47.1%; Score 706; DB 19; Length 1092;  
Best Local Similarity 79.3%; Pred. No. 4.6e-209;  
Matches 864; Conservative 0; Mismatches 220; Indels 5; Gaps 2;

QY 375 GAAATAATGAATGTCAGAGGAAAGTAATCTGTGTGATGTCATGTCACACCGTGTG 434  
Db 6 gaaaaaLaatgaatgTcaaggaagagtggtctctcaatgctgctgctcaactgaat 65  
QY 435 TGTCTGTGTTTGGGAATATGTCACAGAAATTCAGAGGTTGGTGAGAACAGATGGCAGAA 494  
Db 66 ggctgctgtttgggaatacatcaacagaaaccagaaagtgg---cagcagtgctcagag 122  
QY 495 GGACTGCTGTTCCCAAGCTGGTGTAAAAATGGAGCCACACAGTTATCAAGAGACAAACGT 554  
Db 123 ggagctggtgttccagagctggtttaacaatggagactcaacagttacacagaagaaga 182  
QY 555 AGAAGACGCGAGAGAAAAGG---GTAGAAATGGAGATCCGATTTGAAGACCCCTCAGCTATG 611  
Db 183 cgctaLaggaacagaaagaaagaaagaaagaaagagagagagagcttccactagt 242  
QY 612 CGACTGGTTCAATCCAAAACCGCCGGATGTTTTCAGACATGACCCCGTGGAGCGGCC 671  
XX

Db 243 ggaactgggttaatactctgagaaacgcccagaggtcgtgaccaaaacagatggaaggtccc 302  
QY 672 GATTGTGTGGGAAGGACACTTATGACACACAGCTCTGCTGGAAAAAGTACTACGCCACACAGAA 731  
Db 303 agtggatagggaagcactttacaacagacccgtcttagataattattatgccaacagaa 362  
QY 732 ACTCACTGTGGGGCTGACAGATGTTTGCTGTGGAAAGTACATTGAGCATTTACTTTAGAGA 791  
Db 363 aattaccgtgggttgacgggttcttgctgctggaagatacatattgagcattacttggagga 422  
QY 792 CTTTCTGGAGTCTGCTGACATGACTTCACTGTTGGCCATCGGGTCAATTTTACCTCAT 851  
Db 423 gttcttaatactgcaaaatacattctatggttggcccaagtcattcttttaccatcat 482  
QY 852 GATAGACGACACTCCCGGATGCTGCTGACACCTGGAACCTCTTACATTCTCTTACAAGT 911  
Db 483 ggtggatgatatctccagagatgacctttgtagagctgggtcctctcgctccttttaagt 542  
QY 912 CTTTGTAGATCAGGTCTGGAAGAGGTGGCAGGATATCAGCATGATCGGATGAAGACCAT 971  
Db 543 gtttgagatcaagtccgagagaggttggaagacatacagcatgatacgcatgaagacat 602  
QY 972 TGGGGAGCACATCCTGGCCACATCCAGCACGAGGTCGACTTCTCTTCTGTCATGGACCT 1031  
Db 603 cggggagacatactctggccacatccagcaaggtggagcttctctctctgcatggagct 662  
QY 1032 GGATCAAGTCTTTCAAGACAACTTCCGGGTGGAAACTCTGGGCCAGCTGTAGCAGACGT 1091  
Db 663 gatatcaggtcttccaaaacaaactttgggtggagacctgggcccagtcggtgctcagct 722  
QY 1092 CGAGGCTCGGTGTCAAGGCCAGTCCCGAGAGTTTCACTATGAGAGCGGGGAAGTGTCT 1151  
Db 723 acaggcctggtgtacaagggcacatctcagcaggttccacctacgagggcgggaagagtc 782  
QY 1152 GCGCGGTACATCTCCAGAGGGGATTTTACTACCGCGGCCATTTTTCGGAGG 1211  
Db 783 cgcagcctacattccgtttggccaggggagtttttattaccacgcagcatttttgggg 842  
QY 1212 AACGCTACTCACTTCTCAACCTCACAGGGAGTGTCTTTAAGGGATCCTCCAGGACAA 1271  
Db 843 aaacccactcagggttctaaacatacactcaggaagtgcttcaagggaatcctccaggacaa 902  
QY 1272 GAACATGATACATAGAGCCCGGATGATGAGAGCCACCTCAACAAATACATTTCTCTTT 1331  
Db 903 ggaataatgacatagagcagagtggtgcatgataagacccatctaaacaagtatttctct 962  
QY 1332 CAACAAACCCACTAAATCTATCTCCAGAGTATTCGTGGGACTATCAGATAGCGCTGCC 1391  
Db 963 caacaaacccactaaatactcttatccccagaataactgctggtggaatacatataggcatgctc 1022  
QY 1392 TTTCAAGTATTAAGGTGTCAAGGTAGCTTTGGCAGACAAAGAGTATTAATTTGGTTAGAA 1451  
Db 1023 tggatattaggattgtcaagatagcttggcagaaaaaagagataaatttggttagaaa 1082  
QY 1452 TAATGTCTGA 1461  
Db 1083 taacatctga 1092

RESULT 11  
AAT04522  
ID AAT04522 standard; cDNA; 1131 BP.  
XX  
AC AAT04522;  
XX  
DT 01-MAY-1996 (first entry)  
XX  
DE Marmoset alpha-1,3-galactosyltransferase coding region.  
KW Marmoset; alpha-1,3-galactosyltransferase; Immune response; glycoprotein;  
KW alpha-galactosyl epitope; cell membrane; virus; phagocytosis; tumour;  
KW antigen processing; leukemia; lymphoma; myeloma; melanoma; carcinoma;  
KW sarcoma; vaccine; opsonisation; glycoprotein; antibody; anti-Gal; ss.







PR 15-MAY-2000; 2000US-204148P.  
PR 13-JUN-2000; 2000US-0593316.  
PA (GERO-) GERON CORP.

XX Denning C, Clark J;  
XX WPI; 2002-089848/12.  
DR P-PSDB; AAE16319.

XX New ovine tissue devoid of antibody-detectable Gal-alpha(1,3)Gal  
PT determinants, for xenotransplantation, and in the treatment of the  
PT human body by surgery or therapy -  
PS  
PS Claim 18; Page 63-64; 86pp; English.

XX The patent discloses immunologically compatible animal tissue, suitable  
CC for xenotransplantation into human patients. The invention also relates  
CC to an ovine tissue devoid of antibody-detectable Gal-alpha (1,3)Gal (GAL)  
CC determinants which are made by alpha(1,3) galactosyltransferase (GT). The  
CC ovine tissue is useful for treatment of human body by surgery or therapy  
CC and in xenotransplantation, by transplanting the ovine tissue into a  
CC mammal having circulating antibody against Gal alpha(1,3)GT determinants.  
CC Polynucleotide constructs of the invention are useful for inactivating  
CC an alpha1,3 GT gene in an ovine cell. Alpha (1,3) GT sequences are useful  
CC for drug screening and for the production of GAL containing synthetic  
CC oligosaccharides. Sequences of the invention are also useful in gene  
CC therapy. The present sequence is a DNA encoding marmoset alpha (1,3)  
CC GT protein.

XX Sequence 1131 BP; 345 A; 223 C; 283 G; 280 T; 0 other;

Query Match 46.3%; Score 693.8; DB 24; Length 1131;  
Best Local Similarity 78.1%; Pred. No. 3e-205;  
Matches 883; Conservative 0; Mismatches 197; Indels 51; Gaps 2;

QY 382 ATGAATGTCGAAGGAAAGTAATCTGTTGATGCTGATGTCATCCACCGCTGGTTCGGT 441  
DB 1 atgaatgcaaaaggaaagtaattctgtcgatgctggtgtctcaactgtgattgtgtg 60  
QY 442 TTTTGGGAATATGTCACAGA-----ATT 455  
DB 61 ttttgggaatatcatcacagccagagcgtcttcttgttgatatcatcactcaagaag 120  
QY 466 CCAGAGGTTGCTGAGAACAGATGCGAGAGGACTGGTGGTTCACAGCTGTTTAAAAAT 525  
DB 121 ccagaagtgtgatgaacagcagtgtctcagaaggactggtggttctcctggtggtttaacaat 180  
QY 526 GGGACCCACAGTTATCAAGAAGACACAGCTAGAGGACGGAGAGAAAGGGTAG----- 578  
DB 181 gggatccacaattatcaacaaggaggagaagacacagacagacaaagaaaggagaggag 240  
QY 579 -----AAATCGAGATCGCATTTGAAGACCTTCACCTATGGGACTGGTTCATCCAAAG 630  
DB 241 gaacaaaaaaaggagatgacacacagagcttcggtatggaactggtttaatccaaag 300  
QY 631 AACCCGCCGATGTTTACAGTACGACCCCTGGAAGCGCGCGATTTGTGGGAAGGCAC 690  
DB 301 aaacgccagaggttatgacagtgacccaatgaaagcccggtgtgtg99gagggcact 360  
QY 691 TATGACACAGCTCTGCTGGAAAAGTACTACGCCACACAGAAACTCACTGTGGGGCTGACA 750  
DB 361 tacaacaagccatcctctagaaaattattatgccaacagaaaaattaccgtg99ggttgacg 420  
QY 751 GTGTTTCTGCTGGGAAGTACATTGAGCATTACTTACAGACACTTTCTGGAGTCTGCTGAC 810  
DB 421 gttttgcatgtggaagatatattgagcattacttctgagaggagttcgttaacatctgcta 480  
QY 811 ATCTACTTCAATGTTGGCCATCGGTCATATTTTACGTCTATGATAGACACACTCCCGG 870  
DB 481 aggtacttcatggcggccacaaagtcattttatgtcatggcgtgatgctctcaag 540

QY 871 ATGCCTGTCGTGCACCTGAACCCCTACATTCCTTACAAAGTCTTTGAGATCAGGTCGTGAG 930  
DB 541 ggcgcgtttatagagcgtgggtctctcgttctctcaaaagtgtttgaggtccaagcagag 600  
QY 931 AAGAGGTGGCAGGATATCAGCATGATGCGCATGAAGACCATTGGGGAGCACATCCTGGCC 990  
DB 601 aagaggtggcaagacatcagca tga tgcgtatgaagacccatcggggagcaca tcttggcc 660  
QY 991 CACATCAGCACGAGGTGCACTTCTCTCTGCTGATGGACGTGATGAAGTCTTTTCAAGAC 1050  
DB 661 cacatccaacagaggttgacttctctctcgtcgtatgagaccaggtctctccaagac 720  
QY 1051 AACTTCGGGTGGAAACTCTGGGCCAGCTGGTAGCAGACAGCTCCAGGCGCTGGTGGTACAAG 1110  
DB 721 cattttggggtagagaccctggccagtcggtggtctcagctacagggcctggtgtacaag 780  
QY 1111 GCCAGTCCCGAGAAAGTTCACCTATGAGAGCGGGAACCTGTCGGCCGCGTACATTCCATTC 1170  
DB 781 gcagatcctcgtgatgactttacctatgagagcggaagagtcgagcagcatattccattt 840  
QY 1171 GGAGAGGGGGATTTTACTACCAACGCGGCCATTTTGGAGGAACGGCTACTCATTTCTC 1230  
DB 841 ggcagggggattttttattaccatgcagccatttttggaggaacaccgattcaggttctc 900  
QY 1231 AACCTCACCAGGAGGTGCTTTAAGGGGATCCTCCAGGACAAAGAACATGACATAGAGCC 1290  
DB 901 aacatcaccagagtgctttaagggaatcctcctctggaacaagaaatgacatagagcc 960  
QY 1291 CAGTGCATGATCAGAGCCACCTCAACAAATACTTCTTTTCAACAACCCACTAAATC 1350  
DB 961 gagtggcatgatgaagccaccctaaacaagtattctctctcaacaacccctcaaaatc 1020  
QY 1351 CTATCTCCAGAGTATTGCTGGGACTATCAGATAGGCGCTGCTTTCAGATATTTAAAGTGTC 1410  
DB 1021 ttatctccagaaatactgcgtggattatcatatagggccttcagattataaaactgtc 1080  
QY 1411 AAGGTAGCTTGGCAGACAAAGAGTATATTTGGTTAGAAATATATGTCTGA 1461  
DB 1081 aagctatcatggaacaaagagatataatttgggttagaagaagtgtctga 1131

RESULT 13  
AAD26483

ID AAD26483 standard; DNA; 1269 BP.

XX AAD26483;

XX 26-MAR-2002 (first entry)

XX Pig alpha (1, 3) galactosyltransferase DNA.

DE Pig: xenotransplantation; Gal-alpha (1,3)Gal; GAL determinant; surgery;  
KW alpha(1,3) galactosyltransferase; alpha (1, 3) GT; drug screening;  
KW gene therapy; ds.

XX Sus scrofa.

XX Key Location/Qualifiers

FT CDS 16..1131

FT /tag= a

FT /product= "Pig alpha (1, 3) GT protein"

XX WO200188096-A2.

XX 22-NOV-2001.

XX 14-MAY-2001; 2001WO-US15765.

XX 15-MAY-2000; 2000US-204148P.

PR 13-JUN-2000; 2000US-0593316.

XX (GERO-) GERON CORP.

XX

PI Denning C. Clark J;  
XX WPI: 2002-089848/12.  
DR P-PSDB; AAE16320.  
XX  
PT New ovine tissue devoid of antibody-detectable Gal-alpha(1,3)Gal  
PT determinants, for xenotransplantation, and in the treatment of the  
PT human body by surgery or therapy .  
XX  
XX Claim 18; Page 65-66; 86pp; English.  
XX  
CC The patent discloses immunologically compatible animal tissue, suitable  
CC for xenotransplantation into human patients. The invention also relates  
CC to an ovine tissue devoid of antibody-detectable Gal-alpha (1,3)Gal (GAL)  
CC determinants which are made by alpha(1,3) galactosyltransferase (GT). The  
CC ovine tissue is useful for treatment of human body by surgery or therapy  
CC and in xenotransplantation, by transplanting the ovine tissue into a  
CC mammal having circulating antibody against Gal alpha(1,3)GT determinants.  
CC Polynucleotide constructs of the invention are useful for inactivating  
CC an alpha1,3 GT gene in an ovine cell. Alpha (1,3) GT sequences are useful  
CC for drug screening and for the production of GAL containing synthetic  
CC oligosaccharides. Sequences of the invention are also useful in gene  
CC therapy. The present sequence is a DNA encoding pig alpha (1,3) GT  
CC protein.  
XX  
SQ Sequence 1269 BP; 384 A; 259 C; 306 G; 320 T; 0 other;

Query Match 45.68; Score 684.2; DB 24; Length 1269;  
Best Local Similarity 76.68; Pred. No. 3.1e-202;  
Matches 880; Conservative 0; Mismatches 233; Indels 36; Gaps 2;

Qy 371 AGGAGAAATTAATGAATGTCAGGGAAGAAATTAATCCCTGTTGATGCTGATTGCTCAACCG 430  
Db 5 agagaaaaataatgaatgtaacaaagagagtggtctgtaagtctgtctcaactg 64  
Qy 431 TGGTTGCTGCTTTTGGGAAATATGTCACAGAAATTCACAGAGGTT----- 474  
Db 65 taatggtgtgttttgggaatacatcaacagccagaggttcttctgttgataacc 124  
Qy 475 -----GGTCAGAACAGATGGCAGAGGACTGTGTTCCCAAGCTGGT 517  
Db 125 agtcaaaaaaccagaagtttggcagcagtgctcagagagggcgtggttcccgagcgtg 184  
Qy 518 TTAATAATGGGACCCACAGTTTATCAAGAGACACACGTAGAGAGACGGAGAGANAAGG--- 574  
Db 185 ttaacaaatgggactcacagttaccacagaagaagacgtataggaacgaaaggaac 244  
Qy 575 GTAGAAATGGAGATCGCATTTGAAGAGCCTCAGCTATGGGACTGGTTCAATCCAAGAAC 634  
Db 245 aaagaaagaagacaacagaggagagcttcgcgtagtggactggttaatcctcagaaac 304  
Qy 635 GCCCGGATGTTTTCACAGTGAACCCGTTGGAAGCGCCGATTTGCTGGGAGGCACTTATG 694  
Db 305 gccagaggctgtgaccataaccagatggaaaggctccagtggtatgggaaaggcacttaca 364  
Qy 695 ACACAGCTCTGCTGGAAGTACTACGCCACACAGAACTCACTGTGGGGCTGACAGTGT 754  
Db 365 acagagcgtcttagataattattatcccaacagaaaaattaccgtgggcttgacggtt 424  
Qy 755 TTGCTGTGGGAAAGTACATTGAGGATTTACTTAGAAGACTTTTCTGGAGTCTGCTACATGT 814  
Db 425 ttgctgtcggaagatacatatgagatcacttggagggttcttaataatctgcaaatat 484  
Qy 815 ACTTCATGTTGGCATCGGCTGATTTTACGTATGATAGACGACACCTCCCGGATGC 874  
Db 485 acttcaatggcttggccaaagcctatcttttacaacaagggtggaatcctcccgagatgc 544  
Qy 875 CTGCTGTCACCTGAAACCCCTTACATCTTTCATCAAGTCTTTGAGATCAGGCTCTGAGAAGA 934  
Db 545 ctttgatagagctgggtcctctcgttctttaaagtggttggatcaagctccgagaaga 604  
Qy 935 GGTGGCAGGATATCAGCATGATGCGCATGAGACCATTTGGGACACATCTCTGCCCCACA 994

Db 605 ggtggcaagacatcagcatgatgcgcatgaagaccatcggggagcacatccttggccaca 664  
Qy 995 TCCAGCAGGAGTTCGACTTCCTCTCTGCATGGAGTGGATCAAGTCTTTCAAGACAAC 1054  
Db 665 tccagcagaggtggacttctctcttgcataggacgtggatcaggtcttccaaagacagt 724  
Qy 1055 TCGGGGTGGAAACTCTGGGCCAGCTGGTAGCACAGCTCCAGGCCCTGGTGGTACAGGCCA 1114  
Db 725 ttggggtggagaccctggccagtcggtggctcagctacagggcctgggtggacaggcac 784  
Qy 1115 GTCCCGAGAGTTCACCTATGAGAGCGGGGAACCTGTTCGGCCGCTACATTCATTTCGGAG 1174  
Db 785 atcctgacaggttcacactacagagggggaagagtcgcgacgtacattccgtttggcc 844  
Qy 1175 AGGGGATTTTACTTACCACCGCGCCATTTTTGGAGGAACGCCCTACTCACATTTCTCAACC 1234  
Db 845 agggggatctttattaccacgcagccatttttgggggaacacccactcaggttctcaaca 904  
Qy 1235 TCACAGGAGTGTCTTTAAGGGATCTCTCCAGGACAGAAACATGACATAGAACCCCACT 1294  
Db 905 tctctcagaggtgttcaagggaatcctccagggacaaggaatgacatagagccgagt 964  
Qy 1295 GGCATGATGAGAGCCACCTCAACAAATACTTCTTTTCAACAAACCCACTAAAATCTCTAT 1354  
Db 965 ggcatagtgaagccatctaaacaagtatttcttctcaacaacccactaaaaatcttat 1024  
Qy 1355 CTCCAGAGTATTGCTGGGACTATCAGATAGAGCTTCAGATATATTAAGTGTCAAGG 1414  
Db 1025 cccagataactcgtggattatcatacagggatgctctgtggatattaggtattgcaaga 1084  
Qy 1415 TAGCTTGGCAGACAAAAGAGTATAATTTGGTTAGAAATAATGCTCTGACTTCAAAATTTGA 1474  
Db 1085 tagcttggcagaaaaagagataatttgggttagaataaacatctgactcttcaattgtgc 1144  
Qy 1475 TGGAAACTT 1483  
Db 1145 cagcagttt 1153

## RESULT 14

AAT02892  
ID AAT02892 standard; cDNA; 1269 BP.

XX AAT02892;

XX 04-JUN-1996 (first entry)

XX Porcine alpha (1,3) galactosyltransferase cDNA.

XX Transgenic; swine; porcine; alpha (1,3) galactosyltransferase;  
XX antisense; ribozyme; Gal-alpha-1,3-Gal-beta-1-4GlcNAc; epitope;

XX terminal; xenogenic; transplant; rejection; gene therapy; pig; ss.

XX Sus scrofa.

XX Key Location/Qualifiers

FT CDS 16..1131

FT /\*Lag= a

XX W09528412-A1.

XX 26-OCT-1995.

XX 31-MAR-1995; 95WO-US03940.

XX 13-APR-1994; 94US-0228933.

XX (BIOT-) BIOTRANSPLANT INC.

XX (GEO) GEN HOSPITAL CORP.

XX (CHIL-) INST CHILD HEALTH.

XX Baetscher MW, Gustafsson KT, Sachs DH;

XX		WPI; 1995-373759/48.
DR	P-PSDB; AAR85082.	
XX		
PT	Novel transgenic alpha (1.3) galactosyl:transferase negative swine	
PT	- used to produce rejection resistant cells for xenogenic	
PT	transplantation	
XX		
XX	Claim 11; Pages 35-37; 56pp; English.	
PS		
CC	Transgenic swine in which the normal expression of the alpha (1.3)	
CC	galactosyltransferase (ACT) AAR85082 is prevented, are prepd. by	
CC	inhibiting the expression of the ACT gene AT02892 using antisense	
CC	oligonucleotides or ribozyme inactivators in a pluripotent porcine	
CC	embryonic stem cell. It is then inserted into a porcine oocyte	
CC	(from which the pronuclear material has been removed), which is	
CC	itself grown to produce the transgenic swine. Swine which do not	
CC	express ACT will not produce carbohydrate moieties contg. the	
CC	distinctive terminal Gal-alpha-1,3-Gal-beta-1-4GlcNAc epitope,	
CC	which is a significant factor in xenogenic (esp. human) transplant	
CC	rejection of swine grafts. Therefore the swine cells produced in	
CC	the ACT negative transgenic swine are xenogenic transplant	
CC	rejection resistant, and can therefore be used by a transplant	
CC	recipient, or to provide gene therapy.	
XX		
SQ	Sequence 1269 BP; 384 A; 261 C; 304 G; 320 T; 0 other;	
	Query Match	45.4%; Score 681; DB 16; Length 1269;
	Best Local Similarity	76.4%; Pred. No. 3.1e-201;
	Matches 878; Conservative	0; Mismatches 235; Indels 36; Gaps
Oy	371 AGGAGAAATTAATCAATGTCTCAAGGGAAAACGTAATCTCTGTGTGATGCTGATTGCTCTCAACGC 430	
Db		
Db	5 agggaaaaataatgaatgtccaaaggaagtggttctgtccaagtgttgttcacactg 64	
Oy	431 TGGTGTGCGTGTTTGGGAATATGTCTAACAGAANATTCACAGAGTT----- 474	
Db		
Db	65 taatggttggtttgttgggaatatcatcaacgccagaagttcttctgtgatataacc 124	
Oy	475 -----GGTGAGAACAGATGGCAGAAGGACTCGTGGTTCCTCAAGCTGCT 517	
Db		
Db	125 agtcaaaaaccacagaagttggcagcgctcagagggtggtggttcccgagctggt 184	
Oy	518 TTAAAAATGGACCACCCACAGTTATTCAAGACAACAACGTTAGAAAGCACGGAGAGAAAGG--- 574	
Db		
Db	185 ttaacaattggacctcacagattaccacagaagaagacgctaaggcaacagaaagaac 244	
Oy	575 GTAGAAATGGAGATCGCATTTGAAGAGCCTTCAGCTATGGGACTGGTTTCAATCCAAAGAAC 634	
Db		
Db	245 aaagaaaagaagacaaacagagggagagcttcgcgtagtggactggtttaatcctctgagaac 304	
Oy	635 GCCCGGATGTTTTACAGTAGNACCCCGTGGGAAGGCCGATTTGTCGGGAAGGCATTTATG 694	
Db		
Db	305 gccacagagtcgtgaccataaacagaga tgggaaggtcccgatgg catgggaagcacttaca 364	
Oy	695 ACACAGCTCTCTCGGAAAAGTACTACGCCACACAGAAACTCACTGTGGGGCTGCACAGTGT 754	
Db		
Db	365 acagagcgcgtcttaga laattattatgc caaaacagaaaaattac cgtgccttgacycggttt 424	
Oy	755 TTGCTGTGGGAAAGTACATTGAGCATTTACTTAGAAGACTTTTCTGGAGTCTGCTGACATGT 814	
Db		
Db	425 ttgctgtcggaagatacatlgagcattactcttgaggagttctttaa tctgtcaaatcacat 484	
Oy	815 ACTTCATGGTGGCCATCGGGTCATATTTTACGTCTCATGTAGACAGACACCTCCCGGATGC 874	
Db		
Db	485 acttcatggttggccaaaagtcattctttacatcatgtgg gatgatactctccaggtatgc 544	
Oy	875 CTGTCGTGCACCTCAACCCCTCTACATTTCTTACAAGTCTTTGAGATCAGGCTGTGAGAACA 934	
Db		
Db	545 ctltga tagagctgggg lccctctcgctlcccttbaa gtlgtttgaga tcaagtcogagaaga 604	
Oy	935 GGTCGCAGGATATCAGCATGATCGCATGAAGACCACTTTGGGGAGCACATCTTGGCCCCACA 994	



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2002, 11:06:52 ; Search time 25.43 seconds  
(without alignments)  
2266.272 Million cell updates/sec

Title: US-09-863-475A-4  
Perfect score: 2136  
Sequence: 1 MITMLQDLHVKNISMSRSKS.....IKSVKVAQTKYENLVRNV 394

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_17.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2117	99.1	406	11 Q9DBU1	Q9dbul mus musculus
2	1539.5	72.1	376	6 Q28855	Q28855 platyrhini
3	803	37.6	308	6 Q9BH00	Q9bh00 macaca fasc
4	772	36.1	319	11 Q9D4R9	Q9d4r9 mus musculus
5	694.5	32.5	321	6 Q9TTN2	Q9ttn2 pan troglod
6	693	32.4	274	4 Q9UQ65	Q9uq65 homo sapien
7	690.5	32.3	354	4 Q14490	Q14490 homo sapien
8	690	32.3	348	11 Q9ET32	Q9et32 rattus norv
9	689	32.3	274	4 Q9UQ69	Q9uq69 homo sapien
10	689	32.3	274	4 Q9UQ68	Q9uq68 homo sapien
11	689	32.3	274	4 Q9U1R1	Q9u1r1 homo sapien
12	689	32.3	295	4 Q15951	Q15951 homo sapien
13	689	32.3	321	6 Q9RSM2	Q9rsm2 macaca fasc
14	687	32.2	274	4 Q9UQ63	Q9uq63 homo sapien
15	686	32.1	274	4 Q14758	Q14758 homo sapien
16	685	32.1	274	6 Q9TTN5	Q9ttn5 macaca mula
17	684	32.0	274	4 Q9UQ67	Q9uq67 homo sapien
18	684	32.0	274	6 Q9T549	Q9t549 pan troglod
19	683	32.0	274	4 Q9UQ66	Q9uq66 homo sapien

20	683	32.0	274	6	018825	018825 p histo-blo
21	682	31.9	274	6	09TSM3	09tsm3 macaca fasc
22	681	31.9	274	4	015512	015512 h alpha-3-g
23	681	31.9	274	4	09UQ62	09uq62 homo sapien
24	680	31.8	312	6	09TTN4	09ttn4 g glycoprot
25	678.5	31.8	271	6	09TV74	09tv74 pan troglod
26	678	31.7	274	4	09UQ64	09uq64 homo sapien
27	672.5	31.5	332	11	09BQW2	09eqw2 mus musculu
28	671.5	31.4	348	6	09TUD9	09tud9 pan troglod
29	671	31.4	277	6	09TSM1	09tsm1 macaca fasc
30	671	31.4	289	6	09TUE0	09tue0 macaca mula
31	670	31.4	327	11	09EQW3	09eqw3 mus musculu
32	667	31.2	347	4	09UKI5	09uki5 homo sapien
33	664.5	31.1	262	6	09TTN3	09ttn3 pan troglod
34	662	31.0	309	6	09T551	09t551 pan troglod
35	658.5	30.8	364	6	077563	077563 sus scrofa
36	637.5	29.8	347	6	095158	095158 canis famill
37	542.5	25.4	229	4	09C090	09c090 homo sapien
38	540.5	25.3	229	4	09C088	09c088 homo sapien
39	539.5	25.3	229	4	099479	099479 homo sapien
40	539.5	25.3	229	4	09NY02	09ny02 homo sapien
41	538.5	25.2	229	4	09BO97	09bo97 homo sapien
42	537.5	25.2	229	4	09NY01	09ny01 homo sapien
43	516.5	24.2	209	6	09GM53	09gms3 gorilla gor
44	476.5	22.3	190	6	062752	062752 macaca fasc
45	476.5	22.3	190	6	062858	062858 macaca mula

#### ALIGNMENTS

RESULT 1  
Q9DBU1 PRELIMINARY; PRT; 406 AA.  
AC Q9DBU1;  
DT 01-JUN-2001 (TfEMBLrel. 17, Created)  
DT 01-JUN-2001 (TfEMBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TfEMBLrel. 17, Last annotation update)  
DE GLYCOPROTEIN GALACTOSYLTRANSFERASE ALPHA 1, 3.  
GN CGTAL  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=LUNG;  
RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Alizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikola I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki L.,  
RA Hayashizaki Y.;  
RT \*Functional annotation of a full-length mouse cDNA collection.\*;  
RL Nature 409:685-690(2001).  
DR EMBL; AK004753; BAB23534.1; .  
DR MGD; MGI:95704; Ggtal  
SQ SEQUENCE 406 AA; 47866 MW; E08923F47DCD102A CRC64;

```
Query Match 99.1%, Score 2117, DB 11, Length 406;
Best Local Similarity 96.8%, Pred. No. 2.3e-176;
Matches 393; Conservative 1; Mismatches 0; Indels 12; Gaps 1;

Qy 1 MITLQDLHVNKISRSKSETSLPSSSGSGSOEKLNNVKGKILLMLLIVSTVVVFWEYV 60
Db 1 MITLQDLHVNKISRSKSETSLPSSSGSGSOEKLNNVKGKILLMLLIVSTVVVFWEYV 60
Qy 61 N-----RIPEYGENRWKDMWFPSPWFKNGTHSYOEDNVGGRNGRDREE 108
Db 61 NSPEGSFLWIYHTKIPEYGENRWKDMWFPSPWFKNGTHSYOEDNVGGRNGRDREE 120
Qy 109 POLWDFNPKNRDPDLVTPWPKAPVWEGTYDTALLEKYATQKLTGLTVFVAGKYIEH 168
Db 121 POLWDFNPKNRDPDLVTPWPKAPVWEGTYDTALLEKYATQKLTGLTVFVAGKYIEH 180
Qy 169 YLEDFLESADMYFMVGHVIFVYVMDTSTRMPVHLNPLHSLQVFEIRSEKRWQDISMR 228
Db 181 YLEDFLESADMYFMVGHVIFVYVMDTSTRMPVHLNPLHSLQVFEIRSEKRWQDISMR 240
Qy 229 MKTIGELHAIHQHEVDLFCMDVDQVFDQNFVETLGQLVAQLQAWMYKASPEKFTYER 288
Db 241 MKTIGELHAIHQHEVDLFCMDVDQVFDQNFVETLGQLVAQLQAWMYKASPEKFTYER 300
Qy 289 RELSAAIYPFGEGDFYHAAIFGGTPTHILNLTRECFKGILODKKHDIEAQWHDHSLNK 348
Db 301 RELSAAIYPFGEGDFYHAAIFGGTPTHILNLTRECFKGILODKKHDIEAQWHDHSLNK 360
Qy 349 YFLNKPPTKILSPCYWDYQIGLPSDIKSVKVAWQTKYENLVRNV 394
Db 361 YFLNKPPTKILSPCYWDYQIGLPSDIKSVKVAWQTKYENLVRNV 406

RESULT 2
Q28855 PRELIMINARY; PRT; 376 AA.
AC Q28855;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DE ALPHA 1,3 GALACTOSYLTRANSFERASE.
OS Platyrhini (New World monkeys).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates.
OX NCBI_TaxID=9479;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94331837; PubMed=8054718;
RA Henlon T.R., Macher B.A., Anaraki F., Gallili U.;
RT "Defining the minimal size of catalytically active primate alpha 1,3
RT galactosyltransferase: structure-function studies on the recombinant
RT truncated enzyme."
RL Glycobiology 4:193-201(1994).
DR EMBL: S71333; AAB31587.2; -
SQ SEQUENCE 376 AA; 44471 MW; 13C490C0BF2FA2DF CRC64;

Query Match 72.1%, Score 1539.5; DB 6; Length 376;
Best Local Similarity 72.6%, Pred. No. 4.5e-126;
Matches 273; Conservative 43; Mismatches 43; Indels 17; Gaps 2;

Qy 36 MNVKGKIVLLMLIVSTVVVFWEYV-----RIPEYGENRWKDMWFPSPWFKN 83
Db 1 MNVKGKIVLLMLIVSTVVVFWEYVINSPEGSFLWIYHSHKNPEVDSSAQKDMWFPWFNN 60
Qy 84 GTHSYQE-----DNVEGRREKGRNGRDLEPQLWDFNPKNRDPDLVTPWPKAPVWEGT 138
Db 61 GIHNYQDEEDTDKEKEEQQKEDDTTELRLWDFNPKNRDPDLVTPWPKAPVWEGT 120
Qy 139 YDTALLEKYATQKLTGLTVFVAGKYIEHYLEDLFSADMYFMVGHVIFVYVMDTSTR 198
Db 121 YNKAILENYAKQITVGLTVFAIGRVIEHYLEEFVTSANRYFMVGHVIFVYVMDVDSK 180
```

```
Qy 199 MPVHLNPLHSLQVFEIRSEKRWQDISMRMKTIGELHAIHQHEVDLFCMDVDQVFD 258
Db 181 APFIELGLRSLRSEKRVPEKRWQDISMRMKTIGELHAIHQHEVDLFCMDVDQVFD 240
Qy 259 NFGVETLQQLVAQLQAWMYKASPEKFTYERRELSAAYIPFGEGDFYHAAIFGGTPTHIL 318
Db 241 HFGVETLQGSVAQLQAWMYKADPDFTYERRKESAAIYPFGCGDFYHAAIFGGTPIQVL 300
Qy 319 NLTRCFKGILODKKHDIEAQWHDHSLNKYFLNKPPTKILSPCYWDYQIGLPSDIKSV 378
Db 301 NITQCFKGILODKKHDIEAQWHDHSLNKYFLNKPPTKILSPCYWDYHICLPDIDKTV 360
Qy 379 KVAMQTKYENLVRNV 394
Db 361 KLSMQTKYENLVRNV 376

RESULT 3
Q9BH00 PRELIMINARY; PRT; 308 AA.
AC Q9BH00;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE HYPOTHETICAL 36.5 KDA PROTEIN.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FRONTAL CORTEX;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB052526; BAB21880.1; -
KW Hypothetical protein.
SQ SEQUENCE 308 AA; 36456 MW; D7FC82471DD38A3 CRC64;

Query Match 37.6%, Score 803; DB 6; Length 308;
Best Local Similarity 53.1%, Pred. No. 5.3e-62;
Matches 147; Conservative 50; Mismatches 78; Indels 2; Gaps 1;

Qy 105 RIEPQLWDFNPKNRDPDLVTPWPKAPVWEGTYDTALLEKYATQKLTGLTVFVAGVK 164
Db 27 QVEELRLSDWFHPRKRPDVIITKDLAPVWEGTFDQVLEKHRRRITVGLAFFATGR 86
Qy 165 YIEHYLEDLFSADMYFMVGHVIFVYVMDTSTRMPVHLNPLHSLQVFEIRSEKRWQDI 224
Db 87 PAEYELRLFLHSANKHFMGTGRVIFYIMVDAFLQDIPQSPPLTFKFAFEADKRWLEG 146
Qy 225 SMRMKTIGELHAIHQHEVDLFCMDVDQVFDQNFVETLGQLVAQLQAWMYKASPEK 284
Db 147 SLVYKSLGHEHITSHIQDEVDLFSMAVNVQFNEFVETLGPLVAQLHAWMYFRNTKNF 206
Qy 285 TYERRELSAAYIPFGEGDFYHAAIFGGTPTHILNLTRECFKGILODKKHDIEAQWHD 344
Db 207 PYERRPTSAASTPFGGDFYFGSLMVGTPRNILDFTEBYLNGVHTIDIKNGLNSTY--EK 264
Qy 345 HLNKYFLNKPPTKILSPCYWDYQIGLPSDIKSVKVA 381
Db 265 HLNKYFLNKPPTKILSPCYWDYQIGLPSDIKSVKVA 301

RESULT 4
Q9D4R9 PRELIMINARY; PRT; 319 AA.
ID Q9D4R9
AC Q9D4R9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
```

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE 4930568D16RIK PROTEIN.  
 GN 4930568D16RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=TESTIS;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gisi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RT Functional annotation of a full-length mouse cDNA collection.;  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK016248; BAB30163.1; -;  
 DR MGI; MGI:1923109; 4930568D16RIK.  
 SQ SEQUENCE 319 AA; 38038 MW; 519027DC100A3B2A CRC64;

Query Match 36.1%; Score 772; DB 11; Length 319;  
 Best Local Similarity 49.3%; Pred. No. 2.8e-59;  
 Matches 138; Conservative 50; Mismatches 90; Indels 2; Gaps 1;

QY 104 DRIEPPQLDWFNPKRPNVLTVPWKAPIWEGTYDTALLEKYATQKLTGTVFVAVG 163  
 DB 1 DHLEEPHLSTWFDPKRPPDVIATTCGLAPVLWEGTYNREVLEQYKRLNITIGLAVFATG 100  
 QY 164 KYIEHYLEDLESADMYFMVGHVIFVYVMDTDSRMPVHNLPLSHLQVPEIRSEKRWQD 223  
 DB 101 NFSKEPLRRFKSTDKFMVGVYVIFILADSTYNLPEFELGPLRTLKTWRLFEEWCQD 160  
 QY 224 ISMRMKTIGHILAHIOHEVDFLFCMDVQVDFQVDFNFGVETLGLVLAQLQAWMYKASPEK 283  
 DB 161 CNLRNMNMHSHKIIQCVQYEVNLFMAVNTQFNKFNQVETLGLKSAQLHAWMYFKKPRD 220  
 QY 284 FTYERRELSAAYIPFGCEDFYHAAIFGGTPTHLNLTRECFKILQDQKHDIQAQWHD 343  
 DB 221 FPERRTKSAAFIPFEGKDFYHAAIFGGTPTHLNLTRECFKILQDQKHDIQAQWHD 278  
 QY 344 SHLNKYFLFNKPTKILSPYECWDYQIGLPSDIKSVKVAWQ 383  
 DB 279 RHLNKFYFINKPARVLSPEYNWDPRFKTPPEIKHIKIANK 318

RESULT 5  
 ID Q9TTN2 PRELIMINARY; PRT; 321 AA.  
 AC Q9TTN2;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DE 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE ABO HISTO-BLOOD GROUP A TRANSFERASE (FRAGMENT).  
 GN ABO.  
 OS Pan troglodytes (Chimpanzee).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.  
 OX NCBI\_TaxID=9598;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Diamond D.C., Szalay A.A.;  
 RT "ABO Phylogeny";  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF100976; AAF04728.1; -;  
 DR EMBL; AF100972; AAF04728.1; JOINED.  
 DR EMBL; AF100973; AAF04728.1; JOINED.  
 DR EMBL; AF100974; AAF04728.1; JOINED.  
 DR EMBL; AF100975; AAF04728.1; JOINED.  
 KW Transferase.  
 FT NON\_TER 1  
 SQ SEQUENCE 321 AA; 37252 MW; BB8A40843885E7B7 CRC64;

Query Match 32.5%; Score 694.5; DB 6; Length 321;  
 Best Local Similarity 46.1%; Pred. No. 1.6e-52;  
 Matches 143; Conservative 46; Mismatches 108; Indels 13; Gaps 4;

QY 95 GRREKG-----RNGDRIEPPQLDWFNPKN-----RPDLTVTPWKAPIWEGTYDTAL 143  
 DB 12 GSLBERGCMVAREPDHJQHVSLPRMVPQPKVLTPCRKDVLTVPWLAPIWEGTFNIDI 71  
 QY 144 LEKYATQKLTGTVFVAVGKYIEHYLEDLESADMYFMVGHVIFVYVMDTDSRMPVH 203  
 DB 72 LNEQFRQLNTTIGLTVAFAIKYVA-FLKFLTAETAEKHPMVGHRVHYVFTDQPAAPRVT 130  
 QY 204 LNPJLSLOVFIIRSEKRWQDISMMRMTTIGHILAHIOHEVDFLFCMDVQVDFNFGVE 263  
 DB 131 LGTGRQLSVLEVRAYKRWQDVSMRMMEMISDFCQRRLSEVDYLVCDVDMFEFRDHVGE 190  
 QY 264 TLGOLVAQLQAWMYKASPEKPTTYERRELSAAYIPFGCEDFYHAAIFGGTPTHLNLTRE 323  
 DB 191 ILTFLPGLHFGFYGSSGREAFYERRPQSQAYIPRDEGDFYLLGGFPGGSGVQEVORLTRA 250  
 QY 324 CFKGILODKKHIDIEAQWHDHSHLNKRYFLFNKPTKILSPYECWDYQ-IGLPSDIKSVKVAW 382  
 DB 251 CHQAMVQDQANGIEAVWHDHSHLNKRYFLFNKPTKILSPYECWDYQIGLPSDIKSVKVAW 310  
 QY 383 QTKENLVNRN 392  
 DB 311 VPKNHQAVRN 320

RESULT 6  
 ID Q9UQ65 PRELIMINARY; PRT; 274 AA.  
 AC Q9UQ65;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE C1S-AB-SPECIFIC ALPHA 1->3 N-ACETYLGALACTOSAMINYLTTRANSFERASE (FRAGMENT).  
 DE ABO.  
 GN ABO.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yamamoto F.;  
 RT "Human histo-blood group ABO gene locus alleles";  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF134428; AAD26580.1; -;  
 DR EMBL; AF134427; AAD26580.1; JOINED.  
 KW Transferase.  
 FT NON\_TER 1  
 FT NON\_TER 274  
 SQ SEQUENCE 274 AA; 31991 MW; E78E627C028E2955 CRC64;

```
Query Match 32.4%; Score 693; DB 4; Length 274;
Best Local Similarity 49.6%; Pred. No. 1.7e-52;
Matches 136; Conservative 43; Mismatches 93; Indels 2; Gaps 2;

QY 120 RPDVLTVPKAPIWEGTYDPALEKYYATOKLVGLTVFAVGKYYIEHYLEDPLESADM 179
DB 1 RKDVLVTPWLAPIWEGTFENIDILNEQFRQNTTIGLVFAIKKYVA-FLKLFLETAEK 59
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 180 YFMVGHVRIFYVMIDTSMRPMVHNLPLHLSQVFEIRSEKRWQDISMMRMKTIGEHILAH 239
DB 60 HFVGHVRIFYVFTDQAAVPRVLTGTGRQLSVLEVRAYKRWQDVSMRMEMISDFCERR 119
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 240 IQHEVDFLFCMDVQVDFONFQVETLGLVAQLQAWKYKASPEKFTYERRELSAAIIPFG 299
DB 120 FLSEVDYLVCDVDMFEFRDHVGVETLTLPLFGLHPFGYSSREAFYERRRPOQYAI 179
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 300 EGDFFYHAAIEGGTPTHTLNLTRCFKGLQDOKKHIDIEAOHWDESHLNKYFLFNKPTKIL 359
DB 180 EGDFFYLGAFGGVQVQVRLTRACHQAMVVDQANGIEAVWDESHLNKYLRLRHKPTKYL 239
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 360 SPEYCWQYQ-ICLPSDIKSVKAMQTKYENLVN 392
DB 240 SPEYLDQQLLGGWPAVLKRLRTAVPKNHQAVRN 273
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 7
Q14490 PRELIMINARY; PRT; 354 AA.
AC Q14490;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
DE HISTO-BLOOD GROUP A TRANSFERASE.
GN ABO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95118355; PubMed=7598760;
RA Bennett E.P., Steffensen R., Clausen H., Weghuis D.O.,
RA Geurts van Kessel A.;
RT "Genomic cloning of the human histo-blood group ABO locus.";
RL Biochem. Biophys. Res. Commun. 206:318-325(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95298046; PubMed=7779106;
RA Bennett E.P., Steffensen R., Clausen H., Weghuis D.O.,
RA Geurts van Kessel A.;
RT "Genomic cloning of the human histo-blood group ABO locus.";
RL Biochem. Biophys. Res. Commun. 211:347-347(1995).
RN [3]
RP SEQUENCE FROM N.A.
RA Yamamoto F.;
RT "Human histo-blood group ABO gene locus alleles.";
RL Submitted (Mar-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; X84746; CAA59233.1; -.
DR EMBL; X84747; CAA59233.1; JOINED.
DR EMBL; X84748; CAA59233.1; JOINED.
DR EMBL; X84749; CAA59233.1; JOINED.
DR EMBL; X84750; CAA59233.1; JOINED.
DR EMBL; X84751; CAA59233.1; JOINED.
DR EMBL; X84752; CAA59233.1; JOINED.
DR EMBL; AF134413; AAD26573.1; -.
KW Transferase.
SQ SEQUENCE 354 AA; 40950 MW; C344EF584F177134 CRC64;

Query Match 32.3%; Score 690; DB 4; Length 354;
Best Local Similarity 45.8%; Pred. No. 4.1e-52;
Matches 142; Conservative 47; Mismatches 108; Indels 13; Gaps 4;

QY 95 GRREKG-----RNGDRIEEPQLWDFNPKN-----RPDVLTVTPWKAPIVWEGTYDTAL 143
DB 45 GSLRGFCMAVREPHQRLQVSLPRMVYQPKVLTPCRQDVLVTPWLAPIVWEGTFNIDI 104
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 144 LEKYATATOKLVGLTVFAVGKYYIEHYLEDPLESADMYPWVGHVRYFYVMIDTSMRPMV 203
DB 105 LNEQERLQNTTIGLVFAIKKYVA-FLKLFLETAEKHFVGHVRIFYVFTDQAAVPRVT 163
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 204 LNLPLHSQVFEIRSEKRWQDISMMRMKTIGEHILAHIQHEVDFLFCMDVQVDFONFQV 263
DB 164 LCTGRQLSVLEVRAYKRWQDVSMRMEMISDFCERRFSEVDYLVCDVDMFEFRHVG 223
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 264 TLGQLVAQLQAWKYKASPEKFTYERRELSAAIIPFGEGDYFYHAAIFGGTPTHTLNLTR 323
DB 224 ILTLPFGTLHPFGYSSREAFYERRRPOQYAIIPKDEGDFYVLGGFFGGSVQEVQLTRA 283
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 324 CFKGLQDOKKHIDIEAOHWDESHLNKYFLFNKPTKILSPYCHDYO-ICLPSDIKSVKAM 382
DB 284 CHQAMVVDQANGIEAVWDESHLNKYLRLRHKPTKYLSPYLDQQLLGGWPAVLKRLRETA 343
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 383 QTKYENLVN 392
DB 344 VPKNHQAVRN 353
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 8
Q9ET32 PRELIMINARY; PRT; 348 AA.
AC Q9ET32;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE N-ACETYLGALACTOSAMINYLTRANSFERASE A BLOOD GROUP-LIKE ENZYME.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BDIX; TISSUE=STOMACH;
RA Vaidye B., Marlonneau S., Caillieu-Thomas A.L., Bouhours D.,
RA Le Pendu J.;
RT "Rat N-acetylgalactosaminyltransferase A blood group-like enzyme.";
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF264018; AAF74758.2; -.
KW Transferase.
SQ SEQUENCE 348 AA; 40375 MW; FC5E572B100ED7B3 CRC64;

Query Match 32.3%; Score 690; DB 11; Length 348;
Best Local Similarity 48.2%; Pred. No. 4.4e-52;
Matches 132; Conservative 46; Mismatches 94; Indels 2; Gaps 2;

QY 120 RPDVLTVPKAPIWEGTYDPALEKYYATOKLVGLTVFAVGKYYIEHYLEDPLESADM 179
DB 76 RNDVLVTPWLAPIWEGTFENIDILNEQFRQNTTIGLVFAIKKYV-FLKLFLETAEQ 134
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 180 YFMVGHVRIFYVMIDTSMRPMVHNLPLHLSQVFEIRSEKRWQDISMMRMKTIGESHILAH 239
DB 135 HFVGHVRIFYVFTDRPSDVQVPLGAGKRLVLTVRNTRWQDVSMRMEMISHFSQR 194
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 240 IQHEVDFLFCMDVQVDFONFQVETLGLVAQLQAWKYKASPEKFTYERRELSAAIIPFG 299
DB 195 FQHEVDYLVCDVDMKFSQVGVETLSALFGLHPFGYSSREAFYERRRPOQYAIIPD 254
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 300 EGDFFYHAAIEGGTPTHTLNLTRCFKGLQDOKKHIDIEAOHWDESHLNKYFLFNKPTKIL 359
DB 255 EGDFFYAGGFGSVVVEVHLLTKACHQAMVVDQANGIEAVWDESHLNKYLRLRHKPTKYL 314
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 360 SPEYCWQYQ-ICLPSDIKSVKAMQTKYENLVN 392
DB 315 SPEYWDQQLLGGWPSIMKKLRVAVPKNHQALRN 348
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```



```
RESULT 9
Q9U069 PRELIMINARY; PRT; 274 AA.
AC Q9U069;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE AI-SPECIFIC ALPHA 1->3 N-ACETYLGLACTOSAMINYLTRANSFERASE (FRAGMENT).
GN ABO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamamoto F.;
RT "Human histo-blood group ABO gene locus alleles.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF134418; AAD26575.1; -.
DR EMBL; AF134417; AAD26575.1; JOINED.
KW Transferase.
FT NON_TER 1
FT NON_TER 274
SQ SEQUENCE 274 AA; 31961 MW; 12974A4C48954AB6 CRC64;

Query Match 32.3%; Score 689; DB 4; Length 274;
Best Local Similarity 49.3%; Pred. No. 3.9e-52;
Matches 135; Conservative 43; Mismatches 94; Indels 2; Gaps 2;

QY 120 RPDVLTVPWKAPIVWEGTYDTALLEKYATOKLTGTLVFAVGKYIEHYLEDLFLESADM 179
DB 1 RKDVLVTPWLPAPIVWEGTFNIDILNEQFRLQNTTIGLTVFAIKKYVA-FLKLFLETAEK 59
QY 180 YFMVGHVRVIFYVMIDTSMRPVVHLNPLHLSQVFEIRSEKRWODISMRRMKTIGEHLAH 239
DB 60 HFVVGHRVHYVFTDQAAVPRVTLGTGRQLSVLEVRAVKRWODVSMRRMEMISDFCERR 119
QY 240 IQHEVDFLFCMDVDQVDFQDNFVETLGQLVAQLQAWMYKASPEKFTYERRELSAAVYIPFG 299
DB 120 FLSEVDYLCVDVDMFDRHVGVEILTPLFGTLHPGFYGSREAFYERRPQSOAYIPKD 179
QY 300 EGDFFYYHAALFGGTPTTHILNLTRECFKGLQDKKKHDEIAQWHDHSHLNKYFLFNKPTKIL 359
DB 180 EGDFFYLLGGFFGSGVQEVORLTRACHQAMVDAQANGIEAVWHDHSHLNKYLRLRHKPTKVL 239
QY 360 SPEYCWQDYQ-IGLPSDIKSVKVAQTKYENLVN 392
DB 240 SPEYLDQQLLGPWPAVLRLKRLFTAVPKNHQAVRN 273

RESULT 11
Q9UIR1 PRELIMINARY; PRT; 274 AA.
AC Q9UIR1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ABO GLYCOSYLTRANSFERASE (FRAGMENT).
GN ABO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Yip S.P.;
RT "Single-tube multiplex PCR-SSCP analysis distinguishes seven common
RT ABO alleles and readily identifies new alleles.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF182746; AAF07060.1; -.
DR EMBL; AF182745; AAF07060.1; JOINED.
KW Transferase.
FT NON_TER 1
FT NON_TER 274
SQ SEQUENCE 274 AA; 32020 MW; 11F74A4C48954AB6 CRC64;

Query Match 32.3%; Score 689; DB 4; Length 274;
Best Local Similarity 49.3%; Pred. No. 3.9e-52;
Matches 135; Conservative 43; Mismatches 94; Indels 2; Gaps 2;

QY 120 RPDVLTVPWKAPIVWEGTYDTALLEKYATOKLTGTLVFAVGKYIEHYLEDLFLESADM 179
DB 1 RKDVLVTPWLPAPIVWEGTFNIDILNEQFRLQNTTIGLTVFAIKKYVA-FLKLFLETAEK 59
QY 180 YFMVGHVRVIFYVMIDTSMRPVVHLNPLHLSQVFEIRSEKRWODISMRRMKTIGEHLAH 239
DB 60 HFVVGHRVHYVFTDQAAVPRVTLGTGRQLSVLEVRAVKRWODVSMRRMEMISDFCERR 119
QY 240 IQHEVDFLFCMDVDQVDFQDNFVETLGQLVAQLQAWMYKASPEKFTYERRELSAAVYIPFG 299
DB 120 FLSEVDYLCVDVDMFDRHVGVEILTPLFGTLHPGFYGSREAFYERRPQSOAYIPKD 179
```

```
FT NON_TER 1
FT NON_TER 274
SQ SEQUENCE 274 AA; 31977 MW; 71BE047A648E2D8A CRC64;

Query Match 32.3%; Score 689; DB 4; Length 274;
Best Local Similarity 49.3%; Pred. No. 3.9e-52;
Matches 135; Conservative 43; Mismatches 94; Indels 2; Gaps 2;

QY 120 RPDVLTVPWKAPIVWEGTYDTALLEKYATOKLTGTLVFAVGKYIEHYLEDLFLESADM 179
DB 1 RKDVLVTPWLPAPIVWEGTFNIDILNEQFRLQNTTIGLTVFAIKKYVA-FLKLFLETAEK 59
QY 180 YFMVGHVRVIFYVMIDTSMRPVVHLNPLHLSQVFEIRSEKRWODISMRRMKTIGEHLAH 239
DB 60 HFVVGHRVHYVFTDQAAVPRVTLGTGRQLSVLEVRAVKRWODVSMRRMEMISDFCERR 119
QY 240 IQHEVDFLFCMDVDQVDFQDNFVETLGQLVAQLQAWMYKASPEKFTYERRELSAAVYIPFG 299
DB 120 FLSEVDYLCVDVDMFDRHVGVEILTPLFGTLHPGFYGSREAFYERRPQSOAYIPKD 179
QY 300 EGDFFYYHAALFGGTPTTHILNLTRECFKGLQDKKKHDEIAQWHDHSHLNKYFLFNKPTKIL 359
DB 180 EGDFFYLLGGFFGSGVQEVORLTRACHQAMVDAQANGIEAVWHDHSHLNKYLRLRHKPTKVL 239
QY 360 SPEYCWQDYQ-IGLPSDIKSVKVAQTKYENLVN 392
DB 240 SPEYLDQQLLGPWPAVLRLKRLFTAVPKNHQAVRN 273

RESULT 11
Q9UIR1 PRELIMINARY; PRT; 274 AA.
AC Q9UIR1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ABO GLYCOSYLTRANSFERASE (FRAGMENT).
GN ABO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Yip S.P.;
RT "Single-tube multiplex PCR-SSCP analysis distinguishes seven common
RT ABO alleles and readily identifies new alleles.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF182746; AAF07060.1; -.
DR EMBL; AF182745; AAF07060.1; JOINED.
KW Transferase.
FT NON_TER 1
FT NON_TER 274
SQ SEQUENCE 274 AA; 32020 MW; 11F74A4C48954AB6 CRC64;

Query Match 32.3%; Score 689; DB 4; Length 274;
Best Local Similarity 49.3%; Pred. No. 3.9e-52;
Matches 135; Conservative 43; Mismatches 94; Indels 2; Gaps 2;

QY 120 RPDVLTVPWKAPIVWEGTYDTALLEKYATOKLTGTLVFAVGKYIEHYLEDLFLESADM 179
DB 1 RKDVLVTPWLPAPIVWEGTFNIDILNEQFRLQNTTIGLTVFAIKKYVA-FLKLFLETAEK 59
QY 180 YFMVGHVRVIFYVMIDTSMRPVVHLNPLHLSQVFEIRSEKRWODISMRRMKTIGEHLAH 239
DB 60 HFVVGHRVHYVFTDQAAVPRVTLGTGRQLSVLEVRAVKRWODVSMRRMEMISDFCERR 119
QY 240 IQHEVDFLFCMDVDQVDFQDNFVETLGQLVAQLQAWMYKASPEKFTYERRELSAAVYIPFG 299
DB 120 FLSEVDYLCVDVDMFDRHVGVEILTPLFGTLHPGFYGSREAFYERRPQSOAYIPKD 179
```



DB I RKDV LVVTPWLAPIVWEGTFNIDILNEQFRLQN'T'IGLTVFAIRKYYA-FKLFLFLETAEK 59



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2002, 11:05:07 ; Search time 15.93 seconds  
(without alignments)  
1884.041 Million cell updates/sec

Title: US-09-863-475A-4  
Perfect score: 2136  
Sequence: 1 MITMLQDLHVNKISMSRSKS.....IKSVKVAWQTKYENLVNRNV 394  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2136	100.0	394	2 A34417	alpha-1,3-mannosyl
2	1950	91.3	371	2 I49698	alpha-1,3-galactos
3	1533.5	71.8	376	2 A56480	N-acetyllactosamin
4	1510	70.7	371	2 I46583	alpha-1,3-galactos
5	1476.5	69.1	368	2 A44785	N-acetyllactosamin
6	689	32.3	295	2 PC1120	fucosylgalactose a
7	688	32.2	353	1 A34933	fucosylgalactose a
8	637.5	29.8	347	2 JC6126	globoside alpha-N-
9	477.5	22.4	189	2 PC1168	histo-blood group
10	476.5	22.3	189	2 PC1166	histo-blood group
11	469.5	22.0	189	2 PC1165	fucosylgalactose a
12	467.5	21.9	189	2 PC1173	histo-blood group
13	464.5	21.7	189	2 PC1171	histo-blood group
14	462.5	21.7	189	2 PC1169	histo-blood group
15	462.5	21.7	189	2 PC1170	histo-blood group
16	462.5	21.7	189	2 PC1172	histo-blood group
17	114	5.3	26	2 A39769	N-acetyllactosamin
18	99.5	4.7	301	2 G85752	outer membrane pro
19	98	4.6	481	2 T27435	hypothetical prote
20	97	4.5	569	2 A43317	germ cell-less pro
21	96.5	4.5	301	2 B64881	probable outer mem
22	96	4.5	455	2 T29555	hypothetical prote
23	95.5	4.5	361	2 G64354	hypothetical prote
24	95	4.4	505	2 A32261	agarase (EC 3.2.1.
25	94.5	4.4	1078	2 T30879	dynelin heavy chain
26	93	4.4	244	2 H81905	probable phosphoad
27	91.5	4.3	389	2 B84462	hypothetical prote
28	91.5	4.3	455	2 D75043	serine--trRNA synthet
29	91.5	4.3	460	2 G71117	serine--trRNA ligas

30 90.5 4.2 496 2 A69691 ATP-dependent DNA  
31 90 4.2 385 2 A81158 cystathionine beta  
32 1478 2 T38712 ABC transporter Sp  
33 89.5 4.2 612 2 T18463 hypothetical prote  
34 89 4.2 895 2 F75608 conserved hypotheat  
35 88.5 4.1 460 2 T48061 alpha galactosyltr  
36 88.5 4.1 2896 2 T30939 hemocyanin G-type  
37 88 4.1 393 2 S61659 KTR1 protein - yea  
38 88 4.1 1066 2 T30903 arachidonate 8-lip  
39 87.5 4.1 354 2 E81281 hypothetcal prote  
40 87.5 4.1 1941 2 T30554 ubiquitin-protein  
41 87 4.1 385 2 B81949 cystathionine beta  
42 87 4.1 865 2 D72206 valine--trRNA ligas  
43 87 4.1 1102 2 S55100 hypothetcal prote  
44 87 4.1 1169 2 T30207 dynein heavy chain  
45 86.5 4.0 633 2 F75107 abc transporter, A

ALIGNMENTS

RESULT 1

A34417

alpha-1,3-mannosyl-glycoprotein beta-1,4-N-acetylglucosaminyltransferase (EC 2.4.1.14  
C:Species: Mus musculus (house mouse)  
C:Date: 15-Jun-1990 #sequence\_revision 15-Jun-1990 #text\_change 29-Sep-1999  
C:Accession: A34417  
R:Larsen, R.D.; Rajan, V.P.; Ruff, M.M.; Kukowska-Latallo, J.; Cummings, R.D.; Lowe,  
Proc. Natl. Acad. Sci. U.S.A. 86, 8227-8231, 1989  
A:Title: Isolation of a cDNA encoding a murine UDPgalactose:beta-D-galactosyl-1,4-N-a  
A:Reference number: A34417; MUID:90046769  
A:Accession: A34417  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-394 <LAR>

A:Cross-references: GB:M26925; NID:g193419; PIDN:AAA37657.1; PID:g309242  
C:Superfamily: histo-blood group 1 transferase  
C:Keywords: glycosyltransferase; hexosyltransferase; transmembrane protein

Query Match 100.0%; Score 2136; DB 2; Length 394;  
Best Local Similarity 100.0%; Pred. No. 2.7e-174;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSQEKIMNVKGVILLMLIVSTVVVVFWEYV 60  
Db 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSQEKIMNVKGVILLMLIVSTVVVVFWEYV 60  
Qy 61 NRIPEVGENRWQKDWPFPSFKNGTHSYQEDNVEGRREKGRNGDRIEEPQLMDWPNPKNR 120  
Db 61 NRIPEVGENRWQKDWPFPSFKNGTHSYQEDNVEGRREKGRNGDRIEEPQLMDWPNPKNR 120  
Qy 121 PDVLTVTPWKAPIVWEGTYDTALLEKYATOKLTVGLTVFAVGKYIEHYLEDLFLESADMY 180  
Db 121 PDVLTVTPWKAPIVWEGTYDTALLEKYATOKLTVGLTVFAVGKYIEHYLEDLFLESADMY 180  
Qy 181 FMVGHVRYFYVWIDTSMRPVHNLPLHSLOVFEIRSEKRWQDISMRMKTIGEHILAH 240  
Db 181 FMVGHVRYFYVWIDTSMRPVHNLPLHSLOVFEIRSEKRWQDISMRMKTIGEHILAH 240  
Qy 241 QHEVDFLFCMDVDQVQDNFGVETLGQLVAQLQAWMYKASPEKFTYERRELSAAYIPFGE 300  
Db 241 QHEVDFLFCMDVDQVQDNFGVETLGQLVAQLQAWMYKASPEKFTYERRELSAAYIPFGE 300  
Qy 301 GDFYHHAAIFCGTPTHILNLTRECFCGLQDKKHIDIEAQWIDESHLNLYFLFNKPTKILS 360  
Db 301 GDFYHHAAIFCGTPTHILNLTRECFCGLQDKKHIDIEAQWIDESHLNLYFLFNKPTKILS 360  
Qy 361 PEYCWVDYQIGLPSDIKSVKVAWQTKYENLVNRNV 394  
Db 361 PEYCWVDYQIGLPSDIKSVKVAWQTKYENLVNRNV 394

RESULT 2  
149698  
alpha-1,3-galactosyltransferase - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 29-Sep-1999  
C:Accession: I49698  
R:Joziasse, D.H.; Shaper, N.L.; Kim, D.; van der Eijnden, D.H.; Shaper, J.H.  
J. Biol. Chem. 267, 5534-5541, 1992  
A:Title: Murine alpha-1,3-galactosyltransferase: A single gene locus specifies four isoforms  
A:Reference number: A42117; MUID:92184813  
A:Accession: I49698  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-371 <RES>  
A:Cross-references: GB:M85153; NID:g193563; PIDN:AAA37711.1; PID:g457142  
C:Genetics:  
A:Gene: Ggta-1  
C:Superfamily: histo-blood group 1 transferase

Query Match 91.3%; Score 1950; DB 2; Length 371;  
Best Local Similarity 96.5%; Pred. No. 1.7e-158;  
Matches 358; Conservative 1; Mismatches 0; Indels 12; Gaps 1;  
Qy 36 MNVKGKVLILLMLIVSTVVVWFVEYN-----RIPEVGENRWQKDWPFPSWFKN 83  
Db 1 MNVKGKVLILLMLIVSTVVVWFVEYNPDGSLFWIYHTRKIPEVGENRWQKDWPFPSWFKN 60  
Qy 84 GTHSYQEDNVEGRREKGRNGDRIEPQLWDFNPKNRPDLTVTPWKAPIVWEGYDIAL 143  
Db 61 GTHSYQEDNVEGRREKGRNGDRIEPQLWDFNPKNRPDLTVTPWKAPIVWEGYDIAL 120  
Qy 144 LEKYATQKLTGTVFAVGKYLIEHYLEDLESADMYFMVGHVRVIFVYIMDDTSRMPVYH 203  
Db 121 LEKYATQKLTGTVFAVGKYLIEHYLEDLESADMYFMVGHVRVIFVYIMDDTSRMPVYH 180  
Qy 204 LNLHLSLOVFEIRSEKRWQDISMMRMKTIGEHILAHIQHEVDVDFLCMDVQVDFQNFGE 263  
Db 181 LNLHLSLOVFEIRSEKRWQDISMMRMKTIGEHILAHIQHEVDVDFLCMDVQVDFQNFGE 240  
Qy 264 TLGOLVAQLQAWYKASPEKTYERRELSAAYIPFGEGDFYHYHAAIFGGTPTTHILNLTRE 323  
Db 241 TLGOLVAQLQAWYKASPEKTYERRELSAAYIPFGEGDFYHYHAAIFGGTPTTHILNLTRE 300  
Qy 324 CFKGLQDKKHDIEAQWHDHSHLNKYLFLNKPTKILSPYCWYDQIGLPSDIKSVKVAWQ 383  
Db 301 CFKGLQDKKHDIEAQWHDHSHLNKYLFLNKPTKILSPYCWYDQIGLPSDIKSVKVAWQ 360  
Qy 384 TKEYNLVRNV 394  
Db 361 TKEYNLVRNV 371

RESULT 3  
A56480  
N-acetylglucosaminide alpha-1,3-galactosyltransferase (EC 2.4.1.151) - marmoset  
C:Species: Callithrix sp.  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 29-Sep-1999  
C:Accession: A56480  
R:Henion, T.R.; Macher, B.A.; Anataki, F.; Galili, U.  
Glycobiology 4, 193-201, 1994  
A:Title: Defining the minimal size of catalytically active primate alpha1,3 galactosyltransferase  
A:Reference number: A56480; MUID:94331837  
A:Accession: A56480  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-376 <HEN>  
A:Cross-references: GB:S71333; NID:g558051; PIDN:AAB31587.1; PID:g558052  
A:Note: authors translated the codon GTG for residue 251 as Ser  
C:Superfamily: histo-blood group 1 transferase  
C:Keywords: glycosyltransferase; hexosyltransferase; transmembrane protein

Query Match 71.8%; Score 1533.5; DB 2; Length 376;  
Best Local Similarity 72.3%; Pred. No. 5.5e-123;  
Matches 272; Conservative 43; Mismatches 44; Indels 17; Gaps 2;  
Qy 36 MNVKGKVLILLMLIVSTVVVWFVEYN-----RIPEVGENRWQKDWPFPSWFKN 83  
Db 1 MNVKGKVLILLMLIVSTVVVWFVEYNINSPGSLFWIYHSHKNPEVDSSAOKDWPFQMFNN 60  
Qy 84 GTHSYQEDNVEGRREKGRNGDRIEPQLWDFNPKNRPDLTVTPWKAPIVWEGT 138  
Db 61 GTHSYQEDNVEGRREKGRNGDRIEPQLWDFNPKNRPDLTVTPWKAPIVWEGT 120  
Qy 139 YDTALLEKYATQKLTGTVFAVGKYLIEHYLEDLESADMYFMVGHVRVIFVYIMDDTSR 198  
Db 121 YKAILNLYAKOKITVGLTVFAIGRYIEHYLEEFVTSANRYFMVGHKVFYFMVMDVSK 180  
Qy 199 MPVVHLNPLHLSLOVFEIRSEKRWQDISMMRMKTIGEHILAHIQHEVDVDFLCMDVQVDFOD 258  
Db 181 APFIELGPLRSFKVEFKPEKRWQDISMMRMKTIGEHILAHIQHEVDVDFLCMDVQVDFOD 240  
Qy 259 NFGVETLQOLVAQLQAWYKASPEKTYERRELSAAYIPFGEGDFYHYHAAIFGGTPTTHIL 318  
Db 241 HFGVETLQSSAQLQAWYKADPDQDFTYERRKESAAIPFGOGDFYHYHAAIFGGTPTQVL 300  
Qy 319 NUTRECFIGILODKKHDIEAQWHDHSHLNKYLFLNKPTKILSPYCWYDQIGLPSDIKSV 378  
Db 301 NITQCFKGLDKKNDIEAEWHDHSHLNKYLFLNKPDKILSPYCWYDQIGLPSDIKTV 360  
Qy 379 KVAWOTKEYNLVRNV 394  
Db 361 KUSWOTKEYNLVRNV 376

RESULT 4  
I46583  
alpha-1,3-galactosyltransferase - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 29-Sep-1999  
C:Accession: I46583  
R:Strahan, K.M.; Gu, F.; Preece, A.F.; Gustavsson, L.; Andersson, K.  
Immunogenetics 41, 101-105, 1995  
A:Title: cDNA sequence and chromosome localization of pig alpha 1,3 galactosyltransferase  
A:Reference number: I46583; MUID:95104914  
A:Accession: I46583  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-371 <STR>  
A:Cross-references: GB:L36152; NID:g642635; PIDN:AAA73558.1; PID:g642636  
C:Genetics:  
A:Gene: GGTA1  
C:Superfamily: histo-blood group 1 transferase

Query Match 70.7%; Score 1510; DB 2; Length 371;  
Best Local Similarity 72.7%; Pred. No. 5.4e-121;  
Matches 271; Conservative 48; Mismatches 38; Indels 16; Gaps 4;  
Qy 36 MNVKGKVLILLMLIVSTVVVWFVEYN-----RIPEVGENRWQKDWPFPSWFKN 83  
Db 1 MNVKGKVLILLMLIVSTVVVWFVEYNINSPGSLFWIYQSKNPEVGSSA-QRGWPFPSWFKN 59  
Qy 84 GTHSY-QEDNVEGRREKGRNGDRIEPQLWDFNPKNRPDLTVTPWKAPIVWEGTYDT 141  
Db 60 GTHSYHEEEDAIQNEKQKEDNKGELP-LVDWDFNPKRPEVTVTPWKAPIVWEGTYNR 118  
Qy 142 ALLEKYATQKLTGTVFAVGKYLIEHYLEDLESADMYFMVGHVRVIFVYIMDDTSRMPV 201  
Db 119 AVLDNLYAKOKITVGLTVFAVGRIEYHYLEEFISANTYFMVGHKVFYFIMDDISRMPL 178  
Qy 202 VHLNPLHLSLOVFEIRSEKRWQDISMMRMKTIGEHILAHIQHEVDVDFLCMDVQVDFQNF 261  
Db 179 IELGPLRSFKVEIKSEKRWQDISMMRMKTIGEHILAHIQHEVDVDFLCMDVQVDFQNF 238

Qy 262 VETLGQVLAQLQAWMYKASPEKFTYERRELSAAYIPFEGGDFYYHAAIFGGTPTHLNLT 321  
Db 239 VETLGQVLAQLQAWMYKAHPDEFTYERRKESAAYIPFGQDFYYHAAIFGGTPTQVLNIT 298  
Qy 322 RECFCGTLQDKKHIDEAQWHDHSHLNKFLFNKPTKLTLSPEYCDWYQIGLPSDIKSKVVA 381  
Db 299 OECFCGTLQDKKHIDEAQWHDHSHLNKFLFNKPTKLTLSPEYCDWYHIGMSVDIRIVKIA 358  
Qy 382 WOTKEYNLVRNV 394  
Db 359 WOKKEYNLVRNI 371

RESULT 5  
A44785  
N-acetylglucosaminyl 3-alpha-galactosyltransferase (EC 2.4.1.124) - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 19-Mar-1993 #sequence\_revision 19-Mar-1993 #text\_change 29-Sep-1999  
R:Accession: A44785  
R:Joziasse, D.H.; Shaper, J.H.; Van den Eljnden, D.H.; Van Tunen, A.J.; Shaper, N.L.  
J. Biol. Chem. 264, 14290-14297, 1989  
A:Title: Bovine alpha1->3-galactosyltransferase: Isolation and characterization of a cDN  
A:Reference number: A44785; MUID:89340543  
A:Accession: A44785  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-368 <JOZ>  
C:Cross-references: GB:J04989; NID:g163123; PIDN:AAA30558.1; PID:g163124  
A:Superfamily: histo-blood group 1 transferase  
C:Keywords: glycosyltransferase; hexosyltransferase; transmembrane protein

Query Match 69.1%; Score 1476.5; DB 2; Length 368;  
Best Local Similarity 72.2%; Pred. No. 3.8e-118;  
Matches 268; Conservative 41; Mismatches 47; Indels 15; Gaps 3;

Qy 36 MNYKGVKILLMLIVSTVVVWFVEYV-----NRIPVEGENRWQKDMFPSPFKN 83  
Db 1 MNYKGVKILLSMLVSTVIVVWFVEYIHSPEGLFWINPSPNVEPVGSGSIQKGMWLPWFNN 60  
Qy 84 GTHSYOEDNVGRREKGRNGDRIEEDPOLWDFNPKNRPDVLTVTPWKAPTIVWECTYDTAL 143  
Db 61 GYH-EDGDGNEKEQORNEDE-SKULSDWFPNFRPEVVTWTKAPVWVEGTYNRAV 117  
Qy 144 LEKYATQKLTGLTVFVAVGKYIEHYLEDFLESADMYFMVGRHVFYVMIDDTSRMPVVH 203  
Db 118 LDNYAKQKITVGLTVFVAVGRYIEHYLEELFTSANKHFMVGHVPYFVYVDDYSRMLPIE 177  
Qy 204 LNPFLSHQVFEIRSEKRWQDISMRMKTIGEHLAHIQHEVDFLFCMDVDVQDFQDNFGE 263  
Db 178 LGPLRSFKPKIKPEKRWQDISMRMKTIGEHLAHIQHEVDFLFCMDVDVQDFQDKFGE 237  
Qy 264 TLGQVLAQLQAWMYKASPEKFTYERRELSAAYIPFEGGDFYYHAAIFGGTPTHLNLTRE 323  
Db 238 TLGESVAQLQAWMYKADPNDFTYERRKESAAYIPFEGGDFYYHAAIFGGTPTQVLNITQE 297  
Qy 324 CFPKGLQDKKHIDEAQWHDHSHLNKFLFNKPTKLTLSPEYCDWYQIGLPSDIKSKVVAWQ 383  
Db 298 CFPKGLQDKKHIDEAQWHDHSHLNKFLFNKPTKLTLSPEYCDWYHIGLPADIKLVKMSWQ 357  
Qy 384 TKEYNLVRNV 394  
Db 358 TKEYNLVRNV 368

RESULT 6  
PC1120  
fucosylgalactose alpha-N-acetylglucosaminyltransferase (EC 2.4.1.40) A2 allele [valid  
N:Alternate names: histo-blood group A2 transferase  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 08-Sep-2000  
C:Accession: PC1120  
R:Yamamoto, F.; McNeill, P.D.; Hakomori, S.

Biochem. Biophys. Res. Commun. 187, 366-374, 1992  
A:Title: Human histo-blood group A2 transferase coded by A2 allele, one of the A sub  
the carboxyl terminal.  
A:Reference number: PC1120; MUID:92392351  
A:Accession: PC1120  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-295 <YAM>  
C:Cross-references: GB:S44054; NID:g255066; PIDN:AAB23167.1; PID:g255067  
C:Superfamily: histo-blood group 1 transferase  
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 32.3%; Score 689; DB 2; Length 295;  
Best Local Similarity 49.3%; Pred. No. 3.6e-51;  
Matches 135; Conservative 43; Mismatches 94; Indels 2; Gaps 2;  
Qy 120 RPDVLTVPWKAPTIVWECTYDTALLEKYATQKLTGLTVFVAVGKYIEHYLEDFLESADM 179  
Db 1 RKDVLVTPWLPAPTIVWECTFNIDILNQFRLQNTTIGLTVPFAIKKYA-FLKLFLETAEK 59  
Qy 180 YFVGVHRVIFVYVMIDDTSRMPVVHNLPLHSLQVFEIRSEKRWQDISMRMKTIGEHLAH 239  
Db 60 HFVGVHRVHYVFTDQLAAVPRVTLGTGRQLSVLEVRAYKRWQVDSMRMEMISDFCERR 119  
Qy 240 IQHEVDFLFCMDVDVQDFNFGVETLGLQVLAQLQAWMYKASPEKFTYERRELSAAYIPFG 299  
Db 120 FLSEVDVLCVDVDFRDHVGVEILFTLCTLHPGFYSGSSREAFYERRRQSOAYIPKD 179  
Qy 300 EGDYVYHAAIFGGTPTHLNLTRECFCGTLQDKKHIDEAQWHDHSHLNKFLFNKPTKIL 359  
Db 180 EGDYVYLGFGFGGSGVQEVQRLTRACHQAWMYVDQANGIEAVVHDESHLNKYLRLHKPTKV 239  
Qy 360 SPEYCDWYQ-IGLPSDIKSKVVAWQTKYENLVRN 392  
Db 240 SPEYLDWQQLGLGMPAVLURKURFTAVPRNHQAVRN 273

RESULT 7  
A34933  
fucosylgalactose alpha-N-acetylglucosaminyltransferase (EC 2.4.1.40) A1 allele [val  
N:Alternate names: alpha-3-N-acetylglucosaminyltransferase; blood-group substance A  
e: histo-blood group A glycosyltransferase  
C:Species: Homo sapiens (man)  
C:Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 08-Sep-2000  
C:Accession: A34933; S09593; SI3173; PC1164  
R:Yamamoto, F.; Marken, J.; Tsuji, T.; White, T.; Clausen, H.; Hakomori, S.  
J. Biol. Chem. 265, 1146-1151, 1990  
A:Title: Cloning and characterization of DNA complementary to human UDP-GalNAc: Fucal  
A:Reference number: A34933; MUID:90110098  
A:Accession: A34933  
A:Molecule type: mRNA  
A:Residues: 1-353 <YAM1>  
A:Cross-references: GB:J05175; NID:g340077; PIDN:AAA36792.1; PID:g340078  
R:Yamamoto, F.I.; Clausen, H.; White, T.; Marken, J.; Hakomori, S.I.  
Nature 345, 229-233, 1990  
A:Title: Molecular genetic basis of the histo-blood group ABO system.  
A:Reference number: S09593; MUID:90238543  
A:Accession: S09593  
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual  
A:Molecule type: mRNA  
A:Residues: 1-353 <YAM2>  
R:Navaratnam, N.; Findlay, J.B.C.; Keen, J.N.; Watkins, W.M.  
Biochem. J. 271, 93-98, 1990  
A:Title: Purification, properties and partial amino acid sequence of the blood-group-A  
A:Reference number: SI3173; MUID:91024951  
A:Accession: SI3173  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 'X', 65-73, 'IS', 76-77 <NAV>  
R:Kominato, Y.; McNeill, P.D.; Yamamoto, M.; Russell, M.; Hakomori, S.; Yamamoto, F.  
Biochem. Biophys. Res. Commun. 189, 154-164, 1992  
A:Title: Animal histo-blood group ABO genes.

```
Matches 127: Conservative 53: Mismatches 101: Indels 4;

Qy 105 RIEEPQLMDFNPK---NRP-DVLTVTPWKAPIVWEGTYDATALLEKYYATOKLTGLTGVF 160
      ||||| :|: :||||:||||| :|||: :||: |||||
Db 63 RSPHQ-----PKLLEQRPTELLTLPWLADIVSEGTFNPPELLQIYIYOPLNTIGLTVF 116
      ||||| :|: :||||:||||| :|||: :||: |||||
Qy 161 AVGYIEHYLEDLFLESADMYVMGHRVFYVIMDDTSRMPVVHLNPLHLSLOVFEIRSEKR 220
      ||||| :|: :||||:||||| :|||: :||: |||||
Db 117 AVGY-TREVOHFLESAEQFMQGQYVYVIFTNDPAGIPRVPLPGRLLSIPIQRHSR 175
      ||||| :|: :||||:||||| :|||: :||: |||||
Qy 221 WODISMMRMKTIGEHLAHIQHEVDFLFCMDVDQVFDNFQVETLCOLVAQLOAMWYKAS 280
      ||||| :|: :||||:||||| :|||: :||: |||||
Db 176 WEESTRMETISRHAQRAHREVLYLFCVDVDMVFNPNMGWPETLGDVAAIHGPYAVP 235
      ||||| :|: :||||:||||| :|||: :||: |||||
Qy 281 PEKFTYERRELSSAAYIPCEGDFYHYAATFGCTPHILNLTRCFKGIILODKKHIDEAOW 340
      ||||| :|: :||||:||||| :|||: :||: |||||
Db 236 RQQPFYERRHISTATAVENEGDFYGGAVGGRVAKVEYFTTCCHMAILADKANGIMAAW 295
      ||||| :|: :||||:||||| :|||: :||: |||||
Qy 341 HDESHNLNYFLFNKTKILSPEYCWYDIQIGLPSDIKSVKVAMQTKEYNLVRN 392
      ||||| :|: :||||:||||| :|||: :||: |||||
Db 296 QEESHLNRRFISHKPSKVLSPEYLWDDRKPQPSPSLKLIRFSLDRKATSWLRS 347
      ||||| :|: :||||:||||| :|||: :||: |||||

RESULT 9
PC1166
histo-blood group transferase - gorilla (fragment)
C:Species: Gorilla gorilla (gorilla)
C>Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 07-May-1999
C:Accession: PC1166
R:Kominato, Y.; McNeill, P.D.; Yamamoto, M.; Russell, M.; Hakomori, S.; Yamamoto, F.
Biochem. Biophys. Res. Commun. 189, 154-164, 1992
A>Title: Animal histo-blood group ABO genes.
A:Reference number: PC1164; MUID:93080551
A:Accession: PC1166
A:Molecule type: DNA
A:Residues: 1-189 <KOM>
C:Superfamily: histo-blood group 1 transferase

Query Match 22.4%; Score 477.5; DB 2; Length 189;
Best Local Similarity 50.3%; Pred. No. 2.1e-33;
Matches 95: Conservative 25; Mismatches 68; Indels 1; Gaps 1;

Qy 186 RVIFVYMIDDTSRMPVVHLNPLHLSLOVFEIRSEKRWQDISMRMKTIGEHILAHIQEYD 245
      ||||| :|: :||||:||||| :|||: :||: |||||
Db 1 RVHYVYFTDQAANPVRTLCTGRQLSVLEVRAYKRQDVSMRRMEMISDFCERRFLSEVD 60
      ||||| :|: :||||:||||| :|||: :||: |||||
Qy 246 FLFCMDVDQVFDNFGVETLGCOLVAQLOAMWYKASPEKFTYERRELSSAAYIPPEGDFY 305
      ||||| :|: :||||:||||| :|||: :||: |||||
Db 61 YLVCDVDMEDFRDHVGVEILTPLFCTLHPGFYSGSSREAFTYERRRPSQAYIPKDEGDFY 120
      ||||| :|: :||||:||||| :|||: :||: |||||
Qy 306 HAAIFGGTPTHILNLTRCFKGIILODKKHIDEAQHWDESHLNKYFLFNKPTKILSPEYCW 365
      ||||| :|: :||||:||||| :|||: :||: |||||
Db 121 MGAFPGGSVQEVQLTRACHQAMVYDOANGIEAVWHDESHLNKYLLRHKPTKVLSPLEYLW 180
      ||||| :|: :||||:||||| :|||: :||: |||||
Qy 366 DYQ-IGLPS 373
      ||||| :|: :||||:||||| :|||: :||: |||||
Db 181 DQQLLGWPT 189
      ||||| :|: :||||:||||| :|||: :||: |||||

RESULT 10
PC1166
histo-blood group 1 transferase - chimpanzee (fragment)
N:Alternate names: histo-blood group 2 transferase
C:Species: Pan troglodytes (chimpanzee)
C>Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 07-May-1999
C:Accession: PC1166; PC1167
R:Kominato, Y.; McNeill, P.D.; Yamamoto, M.; Russell, M.; Hakomori, S.; Yamamoto, F.
Biochem. Biophys. Res. Commun. 189, 154-164, 1992
A>Title: Animal histo-blood group ABO genes.
A:Reference number: PC1164; MUID:93080551
A:Accession: PC1166
A:Molecule type: DNA
```



C:Species: Papio sp. (baboon)  
C:Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 07-May-1999  
C:Accession: PC1173

C:Species: Pongo pygmaeus (orangutan)  
C:Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 07-May-1999

C:Accession: PC1169  
R:Kominato, Y.; McNeill, P.D.; Yamamoto, M.; Russell, M.; Hakomori, S.; Yamamoto, F.  
Biochem. Biophys. Res. Commun. 189, 154-164, 1992  
A:Title: Animal histo-blood group ABO genes.  
A:Reference number: PC1164; MUID:93080551  
A:Accession: PC1169  
A:Molecule type: DNA  
A:Residues: 1-189 <ROM>  
C:Superfamily: histo-blood group 1 transferase

Query Match 21.7%; Score 462.5; DB 2; Length 189;  
Best Local Similarity 48.7%; Pred. No. 3.9e-32;  
Matches 92; Conservative 26; Mismatches 70; Indels 1; Gaps 1;  
QY 186 RVIFYVMDTSTRMPVVLNPLHSLQVFEIRSEKRWODISMRRMKTIGEHILAHIQHEVD 245  
DB 1 RVHYVFTDQPAVPRVTLGTGROLSVLGVRAYRRQDVSMRRMEMISDFCERRFLSEVD 60  
QY 246 FLFCMDVDQVFQDNFGVETLQQLVAQLQAMWKASPEKFTYERRELSAAYIPFGEGDFYY 305  
DB 61 YLVCVDVDMERDVGVEILTPLFGTLHPGFGYSTREAFYERRRPSQOAYIPKDEGDFYY 120  
QY 306 HAAIFGCTPHILNLTRECFKGILODKKHIEAOHDESHLNKYFLNKPTKILSPYCW 365  
DB 121 LGGFFGSGVQEVQRLTRTCHQAMVDAQANGIEAVVHDESHLNKYLRLRHKPTKVLSPYWL 180  
QY 366 DYQ-IGLPS 373  
DB 181 DQQLGWPA 189

RESULT 15  
PC1170  
histo-blood group 2 transferase - orangutan (fragment)  
C:Species: Pongo pygmaeus (orangutan)  
C:Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 07-May-1999  
C:Accession: PC1170  
R:Kominato, Y.; McNeill, P.D.; Yamamoto, M.; Russell, M.; Hakomori, S.; Yamamoto, F.  
Biochem. Biophys. Res. Commun. 189, 154-164, 1992  
A:Title: Animal histo-blood group ABO genes.  
A:Reference number: PC1164; MUID:93080551  
A:Accession: PC1170  
A:Molecule type: DNA  
A:Residues: 1-189 <ROM>  
C:Superfamily: histo-blood group 1 transferase

Query Match 21.7%; Score 462.5; DB 2; Length 189;  
Best Local Similarity 48.7%; Pred. No. 3.9e-32;  
Matches 92; Conservative 26; Mismatches 70; Indels 1; Gaps 1;  
QY 186 RVIFYVMDTSTRMPVVLNPLHSLQVFEIRSEKRWODISMRRMKTIGEHILAHIQHEVD 245  
DB 1 RVHYVFTDQPAVPRVTLGTGROLSVLGVRAYRRQDVSMRRMEMISDFCERRFLSEVD 60  
QY 246 FLFCMDVDQVFQDNFGVETLQQLVAQLQAMWKASPEKFTYERRELSAAYIPFGEGDFYY 305  
DB 61 YLVCVDVDMERDVGVEILTPLFGTLHPGFGYSTREAFYERRRPSQOAYIPKDEGDFYY 120  
QY 306 HAAIFGCTPHILNLTRECFKGILODKKHIEAOHDESHLNKYFLNKPTKILSPYCW 365  
DB 121 LGGFFGSGVQEVQRLTRTCHQAMVDAQANGIEAVVHDESHLNKYLRLRHKPTKVLSPYWL 180  
QY 366 DYQ-IGLPS 373  
DB 181 DQQLGWPA 189

Search completed: May 10, 2002, 11:07:29  
Job time: 142 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2002, 11:04:47 ; Search time 12.62 Seconds  
(without alignments)  
702.559 Million cell updates/sec

Title: US-09-863-475A-4  
Perfect score: 2136  
Sequence: 1 MITMLQDLHVNKISMSRSKS.....IKSVKVAWQTKYNLVRNV 394

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2136	100.0	394	1	US-07-914-281-4
2	2136	100.0	394	1	US-08-393-246-4
3	2136	100.0	394	1	US-08-525-058A-4
4	2136	100.0	394	2	US-08-696-731-4
5	2136	100.0	394	4	US-09-042-531-4
6	2136	100.0	394	5	PCT-US91-00899-9
7	1950	91.3	371	2	US-08-378-617A-12
8	1539.5	72.1	376	2	US-08-704-548-2
9	1515	70.9	279	5	PCT-US91-00899-8
10	1510	70.7	371	2	US-08-378-617A-10
11	1476.5	69.1	368	2	US-08-378-617A-11
12	690.5	32.3	354	1	US-07-752-101A-51
13	689	32.3	375	1	US-07-752-101A-52
14	688	32.2	353	1	US-07-752-101A-35
15	674.5	31.6	354	1	US-07-752-101A-38
16	672	31.5	353	1	US-07-752-101A-36
17	671.5	31.4	354	1	US-07-752-101A-39
18	671.5	31.4	354	1	US-07-752-101A-41
19	487	22.8	195	1	US-07-752-101A-68
20	256	12.0	154	1	US-07-752-101A-69
21	97	4.5	40	1	US-07-752-101A-10
22	97	4.5	56	1	US-07-752-101A-16
23	87.5	4.1	1720	2	US-08-477-451-12
24	82.5	3.9	822	3	US-08-941-445A-17
25	82	3.8	469	4	US-08-448-489-12
26	81	3.8	569	4	US-08-961-083-154
27	81	3.8	591	4	US-08-961-083-74

28	80.5	3.8	255	1	US-08-152-019A-36	Sequence 36, Appl
29	80.5	3.8	2213	1	US-08-727-034-3	Sequence 3, Appl1
30	80	3.7	469	3	US-08-704-711A-16	Sequence 16, Appl
31	79	3.7	758	1	US-07-756-250-16	Sequence 16, Appl
32	78.5	3.7	42	1	US-07-752-101A-8	Sequence 8, Appl1
33	77.5	3.6	2214	1	US-08-727-034-7	Sequence 7, Appl1
34	76.5	3.6	597	1	US-08-374-155A-12	Sequence 12, Appl
35	76.5	3.6	597	2	US-08-785-396-12	Sequence 12, Appl
36	75	3.5	327	3	US-09-154-874-8	Sequence 8, Appl1
37	75	3.5	462	4	US-08-068-392-3	Sequence 3, Appl1
38	75	3.5	462	4	US-08-396-988-3	Sequence 3, Appl1
39	75	3.5	485	3	US-08-600-658-7	Sequence 7, Appl1
40	75	3.5	485	4	US-09-170-670-6	Sequence 6, Appl1
41	75	3.5	485	4	US-09-193-068-6	Sequence 6, Appl1
42	75	3.5	485	4	US-09-183-412-6	Sequence 6, Appl1
43	75	3.5	485	4	US-09-354-191A-7	Sequence 7, Appl1
44	75	3.5	2332	1	US-07-864-004B-4	Sequence 4, Appl1
45	75	3.5	2332	1	US-08-251-937A-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1  
US-07-914-281-4  
; Sequence 4, Application US/07914281  
; Patent No. 5324663  
; GENERAL INFORMATION:  
; APPLICANT: LOWE, JOHN B.  
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
; TITLE OF INVENTION: GLYCOPOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICANT NUMBER: US/07/914,281  
FILING DATE: 19920720  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Lavalleye, Jean-Paul M. P.  
REGISTRATION NUMBER: 31,451  
REFERENCE/DOCKET NUMBER: 2363-060-55  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)521-4500  
TELEFAX: (703)486-2347  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 394 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-07-914-281-4

Query Match 100.0%: Score 2136; DB 1; Length 394;  
Best Local Similarity 100.0%: Pred No. 1.4e-214;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MITMLQDLHVNKISMSRSKSETSLPSSRSGSQEKIMNVKGKVLMLLIVSTVVVVFWEYV 60

```
Db 1 MITLQDLHVNKISMSRSKSETSLPSSSGSOEKIMNVKGVILLMLLIVSTVVVFWFV 60
Qy 61 NRIPEVGENRQKQDWPFWSFKNGTHSYQEDNVEGRREKGRNGDRIEPQLQWDFNPKNR 120
Db 61 NRIPEVGENRQKQDWPFWSFKNGTHSYQEDNVEGRREKGRNGDRIEPQLQWDFNPKNR 120
Qy 121 PDVLTVPWKAPIVWEGTYDTALLEKYATQKLTGLTVFAVGKVIHYHLEDFLESADMY 180
Db 121 PDVLTVPWKAPIVWEGTYDTALLEKYATQKLTGLTVFAVGKVIHYHLEDFLESADMY 180
Qy 181 FMVGHRIYFYVMIDDTSRMPVYVHNLPLSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240
Db 181 FMVGHRIYFYVMIDDTSRMPVYVHNLPLSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240
Qy 241 QHEVDLFCMDVDQVFDQNFVETLGOLVAQLQAWMYKASPEKTYERRELSAAIYPGE 300
Db 241 QHEVDLFCMDVDQVFDQNFVETLGOLVAQLQAWMYKASPEKTYERRELSAAIYPGE 300
Qy 301 GDFYHAAIFGGTPTTHILNLTRECFCGILQDKKHHIDIEAQWHDHSHLNKYFLFNKPTKILS 360
Db 301 GDFYHAAIFGGTPTTHILNLTRECFCGILQDKKHHIDIEAQWHDHSHLNKYFLFNKPTKILS 360
Qy 361 PEYCDVQIGLPSDIKSVKVAWQTKYNLVRNV 394
Db 361 PEYCDVQIGLPSDIKSVKVAWQTKYNLVRNV 394

RESULT 2
US-08-393-246-4
; Sequence 4, Application US/08393246
; Patent No. 5595900
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,246
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,433
; FILING DATE: 30-MAR-1994
; APPLICATION NUMBER: US 07/914,281
; FILING DATE: 20-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 394 amino acids
; TYPE: amino acid
```

```
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-393-246-4

Query Match 100.0%; Score 2136; DB 1; Length 394;
Best Local Similarity 100.0%; Pred. No. 1.4e-214;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MITLQDLHVNKISMSRSKSETSLPSSSGSOEKIMNVKGVILLMLLIVSTVVVFWFV 60
Db 1 MITLQDLHVNKISMSRSKSETSLPSSSGSOEKIMNVKGVILLMLLIVSTVVVFWFV 60
Qy 61 NRIPEVGENRQKQDWPFWSFKNGTHSYQEDNVEGRREKGRNGDRIEPQLQWDFNPKNR 120
Db 61 NRIPEVGENRQKQDWPFWSFKNGTHSYQEDNVEGRREKGRNGDRIEPQLQWDFNPKNR 120
Qy 121 PDVLTVPWKAPIVWEGTYDTALLEKYATQKLTGLTVFAVGKVIHYHLEDFLESADMY 180
Db 121 PDVLTVPWKAPIVWEGTYDTALLEKYATQKLTGLTVFAVGKVIHYHLEDFLESADMY 180
Qy 181 FMVGHRIYFYVMIDDTSRMPVYVHNLPLSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240
Db 181 FMVGHRIYFYVMIDDTSRMPVYVHNLPLSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240
Qy 241 QHEVDLFCMDVDQVFDQNFVETLGOLVAQLQAWMYKASPEKTYERRELSAAIYPGE 300
Db 241 QHEVDLFCMDVDQVFDQNFVETLGOLVAQLQAWMYKASPEKTYERRELSAAIYPGE 300
Qy 301 GDFYHAAIFGGTPTTHILNLTRECFCGILQDKKHHIDIEAQWHDHSHLNKYFLFNKPTKILS 360
Db 301 GDFYHAAIFGGTPTTHILNLTRECFCGILQDKKHHIDIEAQWHDHSHLNKYFLFNKPTKILS 360
Qy 361 PEYCDVQIGLPSDIKSVKVAWQTKYNLVRNV 394
Db 361 PEYCDVQIGLPSDIKSVKVAWQTKYNLVRNV 394

RESULT 3
US-08-525-058A-4
; Sequence 4, Application US/08525058A
; Patent No. 5770420
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,058A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
```

INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 394 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-525-058A-4

Query Match 100.0%; Score 2136; DB 1: Length 394;  
Best Local Similarity 100.0%; Pred. No. 1.4e-214;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSQEKINNKGKIVLLMLIVSTVVVFEWYV 60  
Db 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSQEKINNKGKIVLLMLIVSTVVVFEWYV 60  
QY 61 NRIPEVGENRWQKDMWFPSPFNKNGTHSYQEDNVNVEGRKGRNCDRIEELPQLDWDFNPKNR 120  
Db 61 NRIPEVGENRWQKDMWFPSPFNKNGTHSYQEDNVNVEGRKGRNCDRIEELPQLDWDFNPKNR 120  
QY 121 PDVLTVPKAPIVWEGTDTALLEKYATQKLTGLTVFVAVGKYIEHYLEDFLESADMY 180  
Db 121 PDVLTVPKAPIVWEGTDTALLEKYATQKLTGLTVFVAVGKYIEHYLEDFLESADMY 180  
QY 181 FMVGHVRVIFYVMIDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHLAHI 240  
Db 181 FMVGHVRVIFYVMIDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHLAHI 240  
QY 241 QHEVDLFCMDVDQVQDNFQVETLQQLVAQLQAMWKASPEKFTYERRELSAAYIPFGE 300  
Db 241 QHEVDLFCMDVDQVQDNFQVETLQQLVAQLQAMWKASPEKFTYERRELSAAYIPFGE 300  
QY 301 GDFYHAAIFGGTPTTHILNLTRECFCGILQDKKHIDEAOWHDESHLNKYLFLNKPTKILS 360  
Db 301 GDFYHAAIFGGTPTTHILNLTRECFCGILQDKKHIDEAOWHDESHLNKYLFLNKPTKILS 360  
QY 361 PEYCWYQIGLPSDIKSVKVMQTKYINLVRNV 394  
Db 361 PEYCWYQIGLPSDIKSVKVMQTKYINLVRNV 394

RESULT 4  
US-08-696-731-4  
Sequence 4, Application US/08696731  
Patent No. 5955347

GENERAL INFORMATION:  
APPLICANT: LOWE, JOHN B.  
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
P.C.  
STREET: 1755 Jefferson Davis Highway, Fourth Floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/696,731  
FILING DATE: 14-AUG-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/393,246  
FILING DATE:

APPLICATION NUMBER: US 08/220,433  
FILING DATE: 30-MAR-1994  
APPLICATION NUMBER: US 07/914,281  
FILING DATE: 20-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Lavalleye, Jean-Paul M. P.  
REGISTRATION NUMBER: 31,451  
REFERENCE/DOCKET NUMBER: 2363-060-55  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)521-4500  
TELEFAX: (703)486-2347  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 394 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-696-731-4

Query Match 100.0%; Score 2136; DB 2: Length 394;  
Best Local Similarity 100.0%; Pred. No. 1.4e-214;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSQEKINNKGKIVLLMLIVSTVVVFEWYV 60  
Db 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSQEKINNKGKIVLLMLIVSTVVVFEWYV 60  
QY 61 NRIPEVGENRWQKDMWFPSPFNKNGTHSYQEDNVNVEGRKGRNCDRIEELPQLDWDFNPKNR 120  
Db 61 NRIPEVGENRWQKDMWFPSPFNKNGTHSYQEDNVNVEGRKGRNCDRIEELPQLDWDFNPKNR 120  
QY 121 PDVLTVPKAPIVWEGTDTALLEKYATQKLTGLTVFVAVGKYIEHYLEDFLESADMY 180  
Db 121 PDVLTVPKAPIVWEGTDTALLEKYATQKLTGLTVFVAVGKYIEHYLEDFLESADMY 180  
QY 181 FMVGHVRVIFYVMIDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHLAHI 240  
Db 181 FMVGHVRVIFYVMIDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHLAHI 240  
QY 241 QHEVDLFCMDVDQVQDNFQVETLQQLVAQLQAMWKASPEKFTYERRELSAAYIPFGE 300  
Db 241 QHEVDLFCMDVDQVQDNFQVETLQQLVAQLQAMWKASPEKFTYERRELSAAYIPFGE 300  
QY 301 GDFYHAAIFGGTPTTHILNLTRECFCGILQDKKHIDEAOWHDESHLNKYLFLNKPTKILS 360  
Db 301 GDFYHAAIFGGTPTTHILNLTRECFCGILQDKKHIDEAOWHDESHLNKYLFLNKPTKILS 360  
QY 361 PEYCWYQIGLPSDIKSVKVMQTKYINLVRNV 394  
Db 361 PEYCWYQIGLPSDIKSVKVMQTKYINLVRNV 394

RESULT 5  
US-09-042-531-4  
Sequence 4, Application US/09042531  
Patent No. 6268193

GENERAL INFORMATION:  
APPLICANT: LOWE, JOHN B.  
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
P.C.  
STREET: 1755 Jefferson Davis Highway, Fourth Floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/042,531  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/393,246  
FILING DATE:  
APPLICATION NUMBER: US 08/220,433  
FILING DATE: 30-MAR-1994  
APPLICATION NUMBER: US 07/914,281  
FILING DATE: 20-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Lavalleye, Jean-Paul M. P.  
REGISTRATION NUMBER: 31,451  
REFERENCE/DOCKET NUMBER: 2363-060-55  
TELEPHONE: (703)521-4500  
TELEFAX: (703)486-2347  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 394 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-09-042-531-4

Query Match 100.0%; Score 2136; DB 4; Length 394;  
Best Local Similarity 100.0%; Pred. No. 1.4e-214;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSQEKIMNVKGVILLMLIVSTVVVFWYEV 60  
DB 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSQEKIMNVKGVILLMLIVSTVVVFWYEV 60

QY 61 NRPEVGENRWQKDWPFPSWFKNGTHSYQEDNVGREGKRGNGRIEPPQLWDWPNKRN 120  
DB 61 NRPEVGENRWQKDWPFPSWFKNGTHSYQEDNVGREGKRGNGRIEPPQLWDWPNKRN 120

QY 121 PDVLTVPKAPIVWEGTYDTALLEKYATQKLTGTVFVAVGKYIEHYLEDLFLESADMY 180  
DB 121 PDVLTVPKAPIVWEGTYDTALLEKYATQKLTGTVFVAVGKYIEHYLEDLFLESADMY 180

QY 181 FMVGHVRVIFYVMIDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
DB 181 FMVGHVRVIFYVMIDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240

QY 241 QHEVDLFCMDVDQVDFQDNFVETLGQLVAQLQAWWYKASPEKFTYERRELSAAYIPGE 300  
DB 241 QHEVDLFCMDVDQVDFQDNFVETLGQLVAQLQAWWYKASPEKFTYERRELSAAYIPGE 300

QY 301 GDFYVHAAIFGGTPTTHLNLTRCFKGILODKKHIDIEAQWHDHSHLNKYFLFNKPTKILS 360  
DB 301 GDFYVHAAIFGGTPTTHLNLTRCFKGILODKKHIDIEAQWHDHSHLNKYFLFNKPTKILS 360

QY 361 PEYCDWYQIGLPSDIKSVKVAWQTKYENLVNRNV 394  
DB 361 PEYCDWYQIGLPSDIKSVKVAWQTKYENLVNRNV 394

RESULT 6  
PCT-US91-00899-9  
Sequence 9, Application PC/TUS9100899  
GENERAL INFORMATION:  
APPLICANT: Lowe, John B.  
TITLE OF INVENTION: Method and Products For the Synthesis of  
TITLE OF INVENTION: Oligosaccharide Structures on Glycoproteins, Glycolipids,

TITLE OF INVENTION: Or as Free Molecules, and For the Isolation of Cloned  
TITLE OF INVENTION: Genetic Sequences That Determine These Structur  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
ADDRESSEE: P.C.  
STREET: 1755 Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/00899  
FILING DATE: 19910214  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Lavalleye Ph.D., Jean-Paul  
REGISTRATION NUMBER: 31,451  
REFERENCE/DOCKET NUMBER: 2363-021-55 PCT  
TELEPHONE: (703)521-5940  
TELEFAX: (703)486-2347  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 394 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
TISSUE TYPE: Blood  
PCT-US91-00899-9

Query Match 100.0%; Score 2136; DB 5; Length 394;  
Best Local Similarity 100.0%; Pred. No. 1.4e-214;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSQEKIMNVKGVILLMLIVSTVVVFWYEV 60  
DB 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSQEKIMNVKGVILLMLIVSTVVVFWYEV 60

QY 61 NRPEVGENRWQKDWPFPSWFKNGTHSYQEDNVGREGKRGNGRIEPPQLWDWPNKRN 120  
DB 61 NRPEVGENRWQKDWPFPSWFKNGTHSYQEDNVGREGKRGNGRIEPPQLWDWPNKRN 120

QY 121 PDVLTVPKAPIVWEGTYDTALLEKYATQKLTGTVFVAVGKYIEHYLEDLFLESADMY 180  
DB 121 PDVLTVPKAPIVWEGTYDTALLEKYATQKLTGTVFVAVGKYIEHYLEDLFLESADMY 180

QY 181 FMVGHVRVIFYVMIDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
DB 181 FMVGHVRVIFYVMIDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240

QY 241 QHEVDLFCMDVDQVDFQDNFVETLGQLVAQLQAWWYKASPEKFTYERRELSAAYIPGE 300  
DB 241 QHEVDLFCMDVDQVDFQDNFVETLGQLVAQLQAWWYKASPEKFTYERRELSAAYIPGE 300

QY 301 GDFYVHAAIFGGTPTTHLNLTRCFKGILODKKHIDIEAQWHDHSHLNKYFLFNKPTKILS 360  
DB 301 GDFYVHAAIFGGTPTTHLNLTRCFKGILODKKHIDIEAQWHDHSHLNKYFLFNKPTKILS 360

QY 361 PEYCDWYQIGLPSDIKSVKVAWQTKYENLVNRNV 394  
DB 361 PEYCDWYQIGLPSDIKSVKVAWQTKYENLVNRNV 394

```

RESULT 7
US-08-378-617A-12
: Sequence 12, Application US/08378617A
: Patent No. 5849991
: GENERAL INFORMATION:
: APPLICANT: d'Apice, Anthony J.F.
: APPLICANT: Pearse, Martin J.
: APPLICANT: Robins, Allan J.
: APPLICANT: Crawford, Robert J.
: APPLICANT: Rathjen, Peter D.
: TITLE OF INVENTION: MATERIALS AND METHODS FOR MANAGEMENT OF
: HYPERACUTE REJECTION IN HUMAN XENOTRANSPLANTATION
: NUMBER OF SEQUENCES: 33
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson
: STREET: 120 South Sixth Street, Suite 2500
: CITY: Minneapolis
: STATE: MN
: COUNTRY: USA
: ZIP: 55402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30B
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/378.617A
: FILING DATE: 26-JAN-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Ellinger, Mark S.
: REGISTRATION NUMBER: 34,812
: REFERENCE/DOCKET NUMBER: 08868/005001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (612) 335-5070
: TELEFAX: (612) 288-9696
: INFORMATION FOR SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 371 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-378-617A-12

```

Query Match	91.3%	Score 1950;	DB 2;	Length 371;
Best Local Similarity	96.5%;	Pred. No. 3.2e-195;		
Matches 359;	Conservative	1;	Mismatches 0;	Indels 12; Gaps 1;
Qy	36	MNVKGVILLMLIVSTVVVVFVEYN-----RPEVGENRWQKWDFPSPFKN	83	
Db				
	1	MNVKGVILLMLIVSTVVVVFMEYVNSPGSFLMIYHTKIPEVGENRWQKWDFPSPFKN	60	
Qy	84	GTHSYQEDNVEGREKGNGDRIEEPQLWDFNEPKNRPDVLTVTPWKAPIVWEGTYTDL	143	
Db				
	61	GTHSYQEDNVEGREKGNGDRIEEPQLWDFNEPKNRPDVLTVTPWKAPIVWEGTYTDL	120	
Qy	144	LEKYATQKLTGVTUTFAVGKYIEHYLEDLFESADMYFMVGHRRVIFYVMIDTSRMPVYH	203	
Db				
	121	LEKYATQKLTGVTUTFAVGKYIEHYLEDLFESADMYFMVGHRRVIFYVMIDTSRMPVYH	180	
Qy	204	LNPJLSLQVFEIRSEKRWODISMMRMKTIGHHILAHIOHEVDLFCMDVOQVFODNFGVE	263	
Db				
	181	LNPJLSLQVFEIRSEKRWODISMMRMKTIGHHILAHIOHEVDLFCMDVOQVFODNFGVE	240	
Qy	264	TLCGLVAQLQAWMKASPEKTYTERRELSAAYIPFEGCDFYYHAHAFGGTPTHILNLTRE	323	
Db				
	241	TLCGLVAQLQAWMKASPEKTYTERRELSAAYIPFEGCDFYYHAHAFGGTPTHILNLTRE	300	
Qy	324	CFKGILODKKHDIQAQNHDSHLNKYELFNKPXTKILSPCYCWDYOIGLPDSDKSVKVAHQ	383	
Db				
	301	CFKGILODKKHDIQAQNHDSHLNKYELFNKPXTKILSPCYCWDYOIGLPDSDKSVKVAHQ	360	

```

Qy 384 TKEYNLVRNV 394
      |||||
Db 361 TKEYNLVRNV 371

RESULT 8
US-08-704-548-2
: Sequence 2, Application US/08704548
: Patent No. 5879675
: GENERAL INFORMATION:
: APPLICANT: GALILI, URI
: APPLICANT: REPIK, PATRICIA M.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR VACCINES
: TITLE OF INVENTION: COMPRISING ALPHA-GALACTOSYL EPITOPES
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.
: STREET: Suite 1800, Two Penn Center Plaza
: CITY: Philadelphia
: STATE: PA
: COUNTRY: USA
: ZIP: 19102
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/704,548
: FILING DATE: 11-SEP-1996
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Monaco, Daniel A.
: REGISTRATION NUMBER: 30,480
: REFERENCE/DOCKET NUMBER: 8760-2 CII
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (215) 568-8383
: TELEFAX: (215) 568-5549
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 376 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-704-548-2

```

Query Match	72.1%	Score	1539.5	DB 2	Length	376	
Best Local Similarity	72.6%	Pred. No.	2.2e-152				
Matches	273	Conservative	43	Mismatches	17	Gaps	2

  

Qy	36	MNVGKVILLMLVSTVVVVVFVEYN-----RIPEVGENRQKQKOWPFPSPFN 83
Db	1	MNVGKVILSLVSTVTVVFWFEYINSPGSFLWYHSKNPEVDSSAQKOWPFGWNN 60
Qy	84	GTHSYQE-----DNVEGRREKGRNDRLEEQQLWDWFPNRPDVLTVTPNKAPIVWEGT 138
Db	61	GIHNYQEEEDTDREKGEERQEKEDDTLEURLMDWFPNKRPEVMTVTQKAPVVWEGT 120
Qy	139	YDTALLEKYATQKLTVGLTVFVAGKYTEHYLEDFLESADMYFMVGHRIYFVYVMTDDTSR 198
Db	121	YNKAILENYAKQKITVGLTVFAIGRYTEHYLEEFVTSANRYFMVGHKVIYFVWVDVSK 180
Qy	199	MPVHLNPLHSQVFEIRSEKRWODISMWRKMTTGEHLAIHQHEVDLFCMDVDQVQFD 258
Db	181	APFTELGPLRSFKFVEYVPEKRWODISMWRKMTTGEHLAIHQHEVDLFCMDVDQVQFD 240
Qy	259	NFGVETLQGLVAQLQAWMYKASPEKTYVERRELSAAYIPFEGGDFYFHAATFGGTPPHIL 318
Db	241	HFGVETLQGSVAQLQAWMYKADPDFTYERKESAYIIPFGGGDFYFHAATFGGTPPIOVL 300
Qy	319	NLTRECFKGILLQDKKHOTEAQWHDESHLUNKYFELFNKPTKILSPCYCWDYOICLPSDIKSV 378





Db 299 QECFKGILQDKNDIEAEMHDESHLNKYLFLNKPTKILSPYCWYHIGMSVDIRVIA 358  
QY 382 WOTKEYNLVRNV 394  
Db 359 WQKKEYNLVRNI 371  
RESULT 11  
US-08-378-617A-11  
: Sequence 11, Application US/08378617A  
: Patent No. 584991  
: GENERAL INFORMATION:  
: APPLICANT: d'Apice, Anthony J.F.  
: APPLICANT: Pearce, Martin J.  
: APPLICANT: Robins, Allan J.  
: APPLICANT: Crawford, Robert J.  
: APPLICANT: Rathjen, Peter D.  
: TITLE OF INVENTION: MATERIALS AND METHODS FOR MANAGEMENT OF  
: TITLE OF INVENTION: HYPERACUTE REJECTION IN HUMAN XENOTRANSPLANTATION  
: NUMBER OF SEQUENCES: 33  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Fish & Richardson  
: STREET: 120 South Sixth Street, Suite 2500  
: CITY: Minneapolis  
: STATE: MN  
: COUNTRY: USA  
: ZIP: 55402  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patent In Release #1.0, Version #1.30B  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/378,617A  
: FILING DATE: 26-JAN-1995  
: CLASSIFICATION: 435  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Ellinger, Mark S.  
: REGISTRATION NUMBER: 34,812  
: REFERENCE/DOCKET NUMBER: 06868/005001  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (612) 335-5070  
: TELEFAX: (612) 288-9696  
: INFORMATION FOR SEQ ID NO: 11:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 368 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
US-08-378-617A-11  
Query Match 69.1%; Score 1476.5; DB 2; Length 368;  
Best Local Similarity 72.2%; Pred. No. 8e-146;  
Matches 268; Conservative 41; Mismatches 47; Indels 15; Gaps 3;  
QY 36 MNVKGKVLMLIVSTVVVFWFVY-----NRIPVEGENRWKQDMWFPWSFKN 83  
Db 1 MNVKGKVLMLVSVIVVFWFVYIHSPECSLEFWINPSRNPVEVGGSIQKGMWLPWFNN 60  
QY 84 GTHSYQEDNVNVEGRREKRGDRIEEPQLMDWPNKRPDVLVTPWKAPIVWEGTYDTAL 143  
Db 61 GYH--BEDGDINEEKQORNEDE-SKLKLSDFNFPFKRPEVVTMTKWKAPVWEGTYNRAV 117  
QY 144 LEKYYATQKLTGLTVFAVGKYTEHYLEDFLESADMYFMVGHVRYFYVVMIDDTSRMPVVH 203  
Db 118 LDNYAKQKLTGLTVFAVGRYTEHYLEEFLTSANKHFMVGHVRYFYVIMVDDYSRMPLE 177  
QY 204 LNPLSQVPEIRSEKRWODISMRMKTIGEHILAHIQHEVDFLFCMDVDVQDFQDNFGVE 263  
Db 178 LGLPSEKFKVPIKPEKRWQDISMRMKTIGEHILVIAHIQHEVDFLFCMDVDVQDFQDNFGVE 237

QY 264 TIGOLVAQLQAMWYKASPEKFTYERRELSAAYIPFGEQDFYHAAIFGGTPTTHILNLTRE 323  
Db 238 TLGESVAQLQAMWYKADPNDFTYERRKESAAYIPFGEQDFYHAAIFGGTPTQVLNITQE 297  
QY 324 CFKGILODKKHIDIEAQWHDHSHLNKYLFLNKPTKILSPYCWYQIGLPSDIKSVKVAWQ 383  
Db 298 CFKGILODKKHIDIEAQWHDHSHLNKYLFLNKPTKILSPYCWYHYHGLPADIKLVKMSWQ 357  
QY 384 TREYNLVRNV 394  
Db 358 TREYNVVRNV 368  
RESULT 12  
US-07-752-101A-51  
: Sequence 51, Application US/07752101A  
: Patent No. 5326857  
: GENERAL INFORMATION:  
: APPLICANT: Yamamoto, Fum1-ichiro  
: APPLICANT: White, Thayer  
: APPLICANT: Hakomori, Sen-itiroh  
: APPLICANT: Clausen, Henrik  
: TITLE OF INVENTION: ABO GENOTYPING  
: NUMBER OF SEQUENCES: 69  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Seed and Berry  
: STREET: 6300 Columbia Center, 701 Fifth Avenue  
: CITY: Seattle  
: STATE: Washington  
: COUNTRY: U.S.  
: ZIP: 98104  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patent In Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/07/752,101A  
: FILING DATE: 19910829  
: CLASSIFICATION: 435  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Sharkey, Richard G.  
: REGISTRATION NUMBER: 32,629  
: REFERENCE/DOCKET NUMBER: 150036.406C1  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 206-622-4900  
: TELEFAX: 206-682-6031  
: TELEX: 3723836  
: INFORMATION FOR SEQ ID NO: 51:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 354 amino acids  
: TYPE: AMINO ACID  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: peptide  
: HYPOTHETICAL: NO  
: ANTI-SENSE: NO  
: FRAGMENT TYPE: N-terminal  
US-07-752-101A-51

Query Match 32.3%; Score 690.5; DB 1; Length 354;  
Best Local Similarity 45.8%; Pred. No. 7.5e-64;  
Matches 142; Conservative 47; Mismatches 108; Indels 13; Gaps 4;

QY 95 GRREKG-----RNGDRIEELQMDWPNPKN-----RPDVLVTPWKAPIVWEGTYDTAL 143  
Db 45 GSLERGFCAVREPRQHLQVSLPRMVPQPKVLTPCRKQVLVTPWLAPIVWEGTFNIDI 104  
QY 144 LEKYYATQKLTGLTVFAVGKYTEHYLEDFLESADMYFMVGHVRYFYVVMIDDTSRMPVVH 203  
Db 105 LNQFRLQNTTIGLTVFAIKKYA-FUKLFLTAERHFVGHVRYFYVFTDQPAVPRVT 163



Qy	95	GREKG-----RNGDRIEPQLWDWENPKN-----RPDLVTWTPMKAPIVWEGTYDTAL	143
Db	45	GSLEFCMAVREPDLHQVSLPRMVPQPKVLTPCRKDVLVVTPMLAPIVWEGTFNIDI	104
Qy	144	LEKYATQKLTUGLTVFVANGKYIEHYLEDFLESADMYFWGVRHVFYVMDITSRMPVWH	203
Db	105	LNQFRLQNTTIGLTVFAIKKYA-FLKULFETAEKFMVGRHVFYVFTDQPAAPRVY	163
Qy	204	LNPLHSLOVEIRSEKRWODISMRMKTTCHEHLAHQHEVDPLFCMDVDVDFODNFGVE	263
Db	164	LTGRQLSVLEVRAYKRWODVSMRRMEMISDFCERFLSEVDVLYLVCVDMFEFRDHVGVE	223
Qy	264	TLQLVAQLQAMWYKASPEKTFVERRELSAAYIPFEGGDFYHAAIFGGTPTHLNLNLTRE	323
Db	224	ILTPLCTLHPGFIGGSSREAFTEYRPSQAYLPKDEGDFYLLGGFPGSGVQVEQLTRTA	283
Qy	324	CFKGIQDKKHIDEAQWHIDESHLNKFYLFNPKPTKILSPCYCWDYO-IGLPSDIKSYKAW	382
Db	284	CHQAMVDAQNGIEAVWHIDESHLNLYLRHKPTKVLSPYLVMDQQLGPNPAVLRLKLFRTA	343
Qy	383	QTKEYNLVRN	392
Db	344	VPKNHOAVRN	353

```

1  GENERAL INFORMATION:
2  APPLICANT: Yamamoto, Fumi-Ichiro
3  APPLICANT: White, Thayer
4  APPLICANT: Hakomori, Sen-Itiroh
5  APPLICANT: Clausen, Henrik
6  TITLE OF INVENTION: ABO GENOTYPING
7  NUMBER OF SEQUENCES: 69
8  CORRESPONDENCE ADDRESS:
9  ADDRESSEE: Seed and Berry
10 STREET: 6300 Columbia Center, 701 Fifth Avenue
11 CITY: Seattle
12 STATE: Washington
13 COUNTRY: U.S.
14 ZIP: 98104
15
16 COMPUTER READABLE FORM:
17 MEDIUM TYPE: Floppy disk
18 COMPUTER: IBM PC Compatible
19 OPERATING SYSTEM: PC-DOS/MS-DOS
20 SOFTWARE: PatentIn Release 1.0, Version #1.25
21
22 CURRENT APPLICATION DATA:
23 APPLICATION NUMBER: US/07/752,101A
24 FILING DATE: 19910829
25 CLASSIFICATION: 435
26 ATTORNEY/AGENT INFORMATION:
27 NAME: Sharkey, Richard G.
28 REGISTRATION NUMBER: 32,529
29 REFERENCE/DOCKET NUMBER: 150036.406C1
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: 206-622-4900
32 TELEFAX: 206-682-6031
33 TELEX: 3723836
34
35 INFORMATION FOR SEQ ID NO: 38:
36 SEQUENCE CHARACTERISTICS:
37 LENGTH: 354 amino acids
38 TYPE: AMINO ACID
39 STRANDEDNESS: single
40 TOPOLOGY: linear
41 MOLECULE TYPE: peptide
42 HYPOTHETICAL: NO
43 ANTI-SENSE: NO
44 FRAGMENT TYPE: N-terminal
45
46 US-07-752-101A-38

```

Query Match 31.6%; Score 674.5; DB 1; Length 354;  
Best Local Similarity 45.2%; Pred. NO. 3.5e-62;  
Matches 140; Conservative 47; Mismatches 110; Indels 13; Gaps 4;



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2002, 11:01:17 ; Search time 25.31 Seconds  
(without alignments)  
1153.097 Million cell updates/sec

Title: US-09-863-475A-4

Perfect score: 2136

Sequence: 1 MITMLQDLHVNKISMSRSK.....IKSVKVAQTKYENLVNRY 394

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_1101.\*  
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2136	100.0	394	12 AAR13750	GDP-Fuc:(beta-D-Ga
2	2136	100.0	394	15 AAR45935	A glycosyltransfer
3	2136	100.0	394	18 AAW13639	Murine alpha(1,3)-
4	1539.5	72.1	376	16 AAR80016	Marmoset alpha-1,3
5	1532	71.7	363	19 AAW49687	Porcine alpha-1,3-
6	1525	71.4	359	15 AAR62508	Galactosyl transfe
7	1525	71.4	359	17 AAR90573	pig alpha(1,3)-gal
8	1513	70.8	375	19 AAW49686	Porcine alpha-1,3-
9	1500	70.2	371	16 AAR85082	Porcine alpha (1,3
10	1493	69.9	371	16 AAR76777	Pig alpha-1,3-gala
11	1431.5	67.0	354	19 AAW49688	Porcine alpha-1,3-

12	1428.5	66.9	342	19 AAW49689	Porcine alpha-1,3-
13	1215	56.9	313	15 AAR62507	Galactosyl transfe
14	690.5	32.3	335	22 AAE05193	Human drug metabol
15	690.5	32.3	354	12 AAR11789	Histo-blood gp. A
16	690.5	32.3	354	15 AAR57020	Human A transferas
17	689	32.3	375	15 AAR57021	Human A transferas
18	688	32.2	353	12 AAR11317	Histo-blood gp. A
19	688	32.2	353	15 AAR57011	Human A transferas
20	687.5	32.2	354	12 AAR11790	Histo-blood gp. B
21	687.5	32.2	354	12 AAR11792	Histo-blood gp. B
22	674.5	31.6	354	15 AAR57013	Human A transferas
23	673.5	31.5	358	15 AAR57014	Human B transferas
24	673	31.5	347	22 AAM39251	Human polypeptide
25	671.5	31.4	354	15 AAR57016	Human B transferas
26	646	30.2	402	22 AAM41037	Human polypeptide
27	487	22.8	195	15 AAR57024	Partial sequence o
28	256	12.0	154	15 AAR57025	hgt4 showing homol
29	230	10.8	149	22 AAM25860	Human protein sequ
30	229	10.7	106	21 AAY86491	Human gene 59-enco
31	216	10.1	100	20 AAY26039	Secreted protein n
32	213	10.0	100	22 AAB73508	Human transferase
33	175	8.2	90	21 AAY86273	Human secreted pro
34	175	8.2	90	21 AAY86323	Human secreted pro
35	175	8.2	90	21 AAY86489	Human gene 59-enco
36	97	4.5	40	15 AAR57004	N-terminal fragmen
37	97	4.5	56	15 AAR57010	N-terminal fragmen
38	97	4.5	432	22 AAG91438	C glutamicum prote
39	97	4.5	432	22 AAB79092	Corynebacterium q1
40	97	4.5	432	22 AAB79725	Corynebacterium q1
41	93.5	4.4	334	22 AAG81594	S. epidermidis ope
42	89.5	4.2	513	14 AAR39900	2184/rhoptry prote
43	88.5	4.1	443	21 AAG13394	Arabidopsis thalia
44	88.5	4.1	460	21 AAG13393	Arabidopsis thalia
45	88.5	4.1	468	21 AAG13392	Arabidopsis thalia

ALIGNMENTS

RESULT 1  
AAR13750  
ID AAR13750 standard; Protein; 394 AA.  
XX  
AC AAR13750;  
DT 07-NOV-1991 (first entry)  
XX  
DE GDP-Fuc:(beta-D-Gal(1,4/1,3))-D-GlcNAc(/Glc)alpha(1,3/1,4)  
DE -fucosyltransferase.  
XX  
KW Glycosyltransferase.  
XX  
OS Mus musculus.  
XX  
PN WO9112340-A.  
XX  
PD 22-AUG-1991.  
XX  
PF 14-FEB-1991; 91WO-US00899.  
XX  
PR 12-DEC-1990; 90US-0627621.  
PR 14-FEB-1990; 90US-0479858.  
PR 14-FEB-1990; 90US-0480133.  
XX  
XX (UNMI ) UNIV OF MICHIGAN.  
PI  
XX Lowc JB;  
XX WPI; 1991-267151/36.  
XX DR N-PSDB; AAQ13331.  
XX  
XX Isolation of gene conveying post-translational characteristic -  
PT e.g. the presence of soluble or membrane bound oligo or

PT polysaccharide or glycosyltransferase.

XX Disclosure; Fig 2; 155pp; English.

CC The amino acid sequence codes for a protein capable of functioning  
 CC as a UDP- Gal: [beta-D-Gal(1,4)]-D-GlcNAc alpha (1,3)galacto-  
 CC syltransferase.. The products of this enzyme, sub-terminal alpha  
 CC (1,3) and alpha(1,4) fucose residues are used in the post-  
 CC translational modification of the oligosaccharides on cell-surface,  
 CC intracellular or secreted proteins or lipids. These can be used for  
 CC the prodn. of diagnostics and therapeutics. There is a single  
 CC transmembrane domain consisting of a 19 amino acid hydrophobic  
 CC segment flanked by basic residues and a large (presumably  
 CC catalytic) C-terminal domain that would ultimately be targeted to  
 CC the lumen of the Golgi. It has two potential N-glycosylation sites  
 CC indicating that as with other glycosyltransferases, it may be  
 CC synthesised as a glycoprotein. It is representative of a Type II  
 CC transmembrane protein. See also AAR13749-R13752.

XX Sequence 394 AA;

Query Match 100.0%; Score 2136; DB 12; Length 394;  
 Best Local Similarity 100.0%; Pred. No. 1e-206;  
 Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSQEKIMNVKGVILLMLIVSTVVVFWEYV 60  
 DB 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSQEKIMNVKGVILLMLIVSTVVVFWEYV 60  
 QY 61 NRPEVCENRWQDWPPSPFKNGTHSYQEDNVEGRKGRNGDRIEPPQLWDFNPKNR 120  
 DB 61 NRPEVCENRWQDWPPSPFKNGTHSYQEDNVEGRKGRNGDRIEPPQLWDFNPKNR 120  
 QY 121 PDVLTVPWKAPIVWEQTYDTALLEKYATQKLTVGTVFAVGKYIEHYLEDLFLESADMY 180  
 DB 121 PDVLTVPWKAPIVWEQTYDTALLEKYATQKLTVGTVFAVGKYIEHYLEDLFLESADMY 180  
 QY 181 FMVGHVRVIFYVMIDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
 DB 181 FMVGHVRVIFYVMIDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
 QY 241 QHEVDLFCHDVOVFDNFGVETLGLVAQLQAWWYKASPEKFTYERRELSAAYIPFGE 300  
 DB 241 QHEVDLFCHDVOVFDNFGVETLGLVAQLQAWWYKASPEKFTYERRELSAAYIPFGE 300  
 QY 301 GDFYHAAIFGGTPTHTLNLTRCFKGILODKKHDIQAOWHDESHLNKYFLFNKPTKILS 360  
 DB 301 GDFYHAAIFGGTPTHTLNLTRCFKGILODKKHDIQAOWHDESHLNKYFLFNKPTKILS 360  
 QY 361 PEYCDWYQIGLPSDIKSVKAWQTKKEYNLVRNV 394  
 DB 361 PEYCDWYQIGLPSDIKSVKAWQTKKEYNLVRNV 394

RESULT 2

AAR45935  
 ID AAR45935 standard; Protein; 394 AA.

XX AAR45935;

XX 26-JUL-1994 (first entry)

DT A glycosyltransferase.

XX Glycosyltransferase; fucosyltransferase; GDP-Fuc; in vitro; cell;  
 XX surface; oligosaccharide.

OS Homo sapiens.

XX WO9402616-A.

XX 03-FEB-1994.

PD

XX 20-JUL-1993; 93WO-US06703.

XX 20-JUL-1992; 92US-0914281.

XX (UNMI ) UNIV MICHIGAN.

XX Lowe JB;

XX WPI; 1994-048874/06.

XX N-PSDB; AAQ56907.

XX DNA fragment encoding a glycosyltransferase - can be used for in  
 PT vitro reactions to modify cell surface oligosaccharides) e.g.  
 PT blood gp. determinants, to protect against transplant rejection

XX Disclosure; Fig 2; 249pp; English.

XX The sequence is that of a human glycosyl transferase. The enzyme  
 CC may be non glycosylated. This prevents premature loss of enzyme  
 CC activity. It can also be used in in vitro reactions to modify cell  
 CC surface oligosaccharide mols. e.g. blood group determinants.  
 CC See also AAR45933-9.

XX Sequence 394 AA;

Query Match 100.0%; Score 2136; DB 15; Length 394;  
 Best Local Similarity 100.0%; Pred. No. 1e-206;  
 Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSQEKIMNVKGVILLMLIVSTVVVFWEYV 60  
 DB 1 MITMLQDLHVNKISMSRSKSETSLPSSRSQSQEKIMNVKGVILLMLIVSTVVVFWEYV 60  
 QY 61 NRPEVCENRWQDWPPSPFKNGTHSYQEDNVEGRKGRNGDRIEPPQLWDFNPKNR 120  
 DB 61 NRPEVCENRWQDWPPSPFKNGTHSYQEDNVEGRKGRNGDRIEPPQLWDFNPKNR 120  
 QY 121 PDVLTVPWKAPIVWEQTYDTALLEKYATQKLTVGTVFAVGKYIEHYLEDLFLESADMY 180  
 DB 121 PDVLTVPWKAPIVWEQTYDTALLEKYATQKLTVGTVFAVGKYIEHYLEDLFLESADMY 180  
 QY 181 FMVGHVRVIFYVMIDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
 DB 181 FMVGHVRVIFYVMIDDTSRMPVHNLPLHSLQVFEIRSEKRWQDISMMRMKTIGEHILAH 240  
 QY 241 QHEVDLFCHDVOVFDNFGVETLGLVAQLQAWWYKASPEKFTYERRELSAAYIPFGE 300  
 DB 241 QHEVDLFCHDVOVFDNFGVETLGLVAQLQAWWYKASPEKFTYERRELSAAYIPFGE 300  
 QY 301 GDFYHAAIFGGTPTHTLNLTRCFKGILODKKHDIQAOWHDESHLNKYFLFNKPTKILS 360  
 DB 301 GDFYHAAIFGGTPTHTLNLTRCFKGILODKKHDIQAOWHDESHLNKYFLFNKPTKILS 360  
 QY 361 PEYCDWYQIGLPSDIKSVKAWQTKKEYNLVRNV 394  
 DB 361 PEYCDWYQIGLPSDIKSVKAWQTKKEYNLVRNV 394

RESULT 3

AAR13639  
 ID AAR13639 standard; Protein; 394 AA.

XX AAR13639;

XX 19-JUN-1997 (first entry)

DT Murine alpha(1,3)-galactosyltransferase.

XX Alpha(1,3)-galactosyltransferase; glycosylation; oligosaccharide.

XX Mus sp.

```

XX PN WO9709421-A1.
XX PD 13-MAR-1997.
XX PF 06-SEP-1996; 96WO-US13816.
XX PR 08-SEP-1995; 95US-0525058.
XX PA (UNMI ) UNIV MICHIGAN.
XX PI Legault DJ, Lowe JB;
XX PWPI; 1997-192897/17.
XX DR N-PSDB; AAT61676.
XX PT New recombinant fucosyltransferase proteins - useful for modifying
XX PT cell surface oligosaccharide structures
XX PS Example 2; Page 272-274; 329pp; English.
XX CC Murine UDP-Gal:beta-D-Gal(1,4)-D-GlcNAc alpha(1,3)-
XX CC galactosyltransferase (AAW13639) catalyses the a transglycosylation
XX CC reaction between UDP-Gal and N-acetyllactosamine and is associated
XX CC with surface-localised expression of Gal(alpha1-3)Gal linkages.
XX CC Its amino acid sequence was deduced from a cDNA clone (AAT61676)
XX CC obt'd. by transfecting COS-1 cells with cDNA derived from mouse F9
XX CC teratocarcinoma cells, and screening the transfected cells for
XX CC surface-localised Gal(alpha1-3)Gal linkages. When expressed in
XX CC animal cell lines, the enzyme provides specific capabilities with
XX CC respect to post-translational modification of the oligosaccharides
XX CC of expressed proteins or lipids. The enzyme can also be used to
XX CC raise antibodies and to screen for inhibitor cpds.
XX CC Sequence 394 AA;
XX SQ
Query Match 100.0%; Score 2136; DB 18; Length 394;
Best Local Similarity 100.0%; Pred. No. 1e-206;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MITMLQDLHVNKLSMSRKSSETSLPSSRSGSOEIKMNVKGVILLMLIVSTVVVFWEYV 60
DB 1 mitmlqdlhvnklsmsrkssetsipsrsgsqekimnvkgvillmlivstvvvfweyv 60
QY 61 NRPEVGENRWQDMWPPSFKNGTHSYQEDNVGRRKGRNGDRIEEPQLWDFWPNKNR 120
DB 61 nrpevgenrwqdmwppsfkngthsyqednvgrrekgrngdrieepqlwdfwnpknr 120
QY 121 PDVLTVPWKAPIWEGTYDTALLEKYATOKLTVGLTVFAVGKYIEHYLEDLESADMY 180
DB 121 pdvltvtpwkapiwegttydtallekyatqktlvtgltvfavgkylehyledflesadmy 180
QY 181 FMVGHRYFVYVIMDDTSRMPVHNLPLHSLQVFEIRSEKRWODISMMRMKTIGEHILAH 240
DB 181 fmvghryfvymddtsrmpvvhnlplhslqvfeirsekrwodismmrmktigehilahi 240
QY 241 QHEVDLFLCNDVQVDFONFVETLGLQVAQLQAWMYKASPEKFTYERRELSAAIIPFGE 300
DB 241 qhevdflfncndvqdvdfonfvettlglqlvaqlqawmykaspekftyerrelsaaiipfge 300
QY 301 GDFYHAAIFGGTPTTHLNLNLTRECFCGILQDKKHDIQAQWHDSEHLNKNYFLFNKPTK 360
DB 301 gdfyhaaifggtpthlnlnltrecfcgilkdkkhdieaqwhdeshlnknlyflfnkptk 360
QY 361 PEYCDWQIGLPSDIKSVKAWQTKENLVNINV 394
DB 361 peycdwqigilpsdiksvkawkqtkenlvninv 394
RESULT 4
AAR80016
ID AAR80016 standard; Protein; 376 AA.

```

```

XX AAR80016;
XX AC 01-MAY-1996 (first entry)
XX DT Harmoset alpha-1,3-galactosyltransferase.
XX DE Marmoset; alpha-1,3-galactosyltransferase; immune response; glycoprotein;
XX KW alpha-galactosyl epitope; cell membrane; virus; phagocytosis; tumour;
XX KW antigen processing; leukemia; lymphoma; myeloma; melanoma; carcinoma;
XX KW sarcoma; vaccine; opsonisation; glycoprotein; antibody; anti-Gal.
XX OS Callithrix jacchus.
XX PN WO9524924-A1.
XX PD 21-SEP-1995.
XX PF 13-MAR-1995; 95WO-US03156.
XX PR 15-MAR-1994; 94US-0213200.
XX CC (UYHA-) UNIV HAHNEMANN & MEDICAL COLLEGE PENNSYL.
XX PI Gallili U, Repik PM;
XX DR WPI; 1995-336816/43.
XX DR N-PSDB; AAT04522.
XX PT Association of an alpha-galactosyl epitope with a tumour or viral
XX PT antigen - is administered to anti-Gal synthesising animals to induce
XX PS an immune response
XX PS Disclosure; Fig 9; 85pp; English.
XX CC The amino acid sequence of the marmoset alpha-1,3-galactosyltransferase.
XX CC The enzyme can be used in methods of enhancing an immune response by
XX CC associating the alpha-galactosyl epitope with a cell membrane or viral
XX CC glycoprotein. The alpha-galactosyl epitope enhances phagocytosis and
XX CC subsequent processing of the antigen. The method is useful in the
XX CC treatment of tumours e.g. leukemia, lymphoma, myeloma, melanoma,
XX CC carcinoma and sarcoma, or for the generation of viral vaccines by
XX CC opsonising a viral glycoprotein. The alpha-galactosyl epitope enhances
XX CC recognition of the antigen in an animal that synthesises the naturally
XX CC occurring antibody - anti-Gal.
XX SQ Sequence 376 AA;
Query Match 72.1%; Score 1539.5; DB 16; Length 376;
Best Local Similarity 72.6%; Pred. No. 1.2e-146;
Matches 273; Conservative 43; Mismatches 43; Indels 17; Gaps 2;
QY 36 MNVKGKVVILLMLIVSTVVVFWEYVN-----RPEVGENRWQDMWPPSPFKN 83
DB 1 mnvkgkvillmlivstvvvfweyinspedsflwyhsknpveddsaqkdwfpfgvnn 60
QY 84 GTHSYQE-----DNVEGRREKGRNGDRIEEPQLWDFWPNKRPDLTVTPWKAPIWEGT 138
DB 61 ghnyqqceedtdkegseeqkddtdtelrlwdfnpkkrpvmvtvtqkavvvwgt 120
QY 139 YDTALLEKYATOKLTVGLTVFAVGKYIEHYLEDLESADMYFMVGHRYFVYVIMDDTSR 198
DB 121 ynkaillenyyakqkltvgtlvtfalgrylehylesfvtsanryfmvghkfyvmvddvsk 180
QY 199 MPVVHNLPLHSLQVFEIRSEKRWODISMMRMKTIGEHILAHQHEVDLFLCNDVQVDF 258
DB 181 apfiegplrsfkfvefkpkewrdqismrmktigehilahlqhevdflfncndvqvfq 240
QY 259 NFGVETLGLQVAQLQAWMYKASPEKFTYERRELSAAIIPFGEQDFYHAAIFGGTPTPH 318
DB 241 hfgvetlgsvaqlqawmykasppddftyrerkesaaylpfggdfyhaaifggtpth 300

```

Qy 319 NLTRCFKGILODKKHDIEAOWHDESHLNKYFLFNKPTKILSPCYCWDYQIGLPSDIKSV 378  
 Db 301 nitgcfkgilldkndleaeahdeshlnkyfllnkpskllspsycwdyhgllpsdiktv 360

Qy 379 KVAMQTKYNLVRNV 394  
 Db 361 klsqctkeynlvrknv 376

RESULT 5  
 ID AAW49687 standard; Protein; 363 AA.  
 XX AAW49687;  
 AC AAW49687;  
 DT 10-NOV-1998 (first entry)  
 XX Porcine alpha-1,3-galactosyl transferase isoform 2.  
 DE Isoform; porcine; enzyme; alpha-1,3-galactosyl transferase; galactose;  
 KW sugar; N-acetylglucosamine; glycoprotein; glycolipid; antibody; pig;  
 KW graft tissue rejection; organ transplantation; xenotransplant.  
 XX Sus scrofa.  
 OS FR2751346-A1.  
 XX 23-JAN-1998.  
 PD 19-JUL-1996; 96FR-0009077.  
 PF 19-JUL-1996; 96FR-0009077.  
 PR (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
 XX Pourcel C, Souillou JP, Vanhove B;  
 PI WPI; 1998-112876/11.  
 XX N-PSDB; AAV49454.  
 DR Transgenic non-human donors of organs for human recipients -  
 PT containing DNA encoding antibodies that inhibit graft rejection  
 XX Claim 4; Page 36-37; 71pp; French.  
 CC This sequence represents isoform 2 of the porcine enzyme  
 CC alpha-1,3-galactosyl transferase (alpha-1,3-GT). The enzyme catalyses  
 CC the attachment of a galactose sugar molecule on the N-acetylglucosamine  
 CC moiety found on surface glycoproteins and glycolipids. These sugar  
 CC molecules are partly responsible for raising anti-graft antibodies, which  
 CC lead to graft tissue rejection. The invention relates to a method of  
 CC inhibiting the graft rejection mechanism by introducing the sequence  
 CC encoding an antibody targeted to alpha-1,3-GT into the cells of animal,  
 CC especially a pig, from whom organs may be used for xenotransplants.  
 CC Neutralisation of the alpha-1,3-GT leads to tissues or organs lacking  
 CC the galactose on the glycoproteins and glycolipids, thus preventing  
 CC induction of the rejection response.  
 XX Sequence 363 AA;

Query Match 71.78; Score 1532; DB 19; Length 363;  
 Best Local Similarity 75.26; Pred. No. 6.7e-146;  
 Matches 273; Conservative 47; Mismatches 39; Indels 4; Gaps 3;

Qy 34 KIMVKGKIVLLMLIVSTVVVFWVEYVNRIPVEGENRWOKDMWFPKNGTHSY--OED 91  
 Db 3 kimvkgrrvllmlivscvmvllvlnrnpvegssa-qrgwwfpswfnngthsyheeed 61

Qy 92 NVEGRKGRNGRDIEEPQLQWDFNPKNRPDVLTVTPWKAPVWEGYDTALLEKYATQ 151  
 Db 62 aignekeqrkedngelp-lvdwnpekrpevltitwkapvvegtynravldnyyak 120

Qy 152 KLTVGLTVFANGKYIEHYLEDPLESADMYFVWGHVRVIFYVMIDTSTRMPVVLNPLHSLQ 211  
 Db 121 kltvgltvlangrylehyleeiflsantymvghkvifylmvddlsrmplllelglprsfk 180

Qy 212 VFEIRSEKRWODISMRMKTICEHILAHIQHEVDLFCMDVDVQDQNFQVETLQOLVAQ 271  
 Db 181 vfeiksekrwqdismmrmtkigeihlahiqhevdficmdvdvfgqnnfgvetigqsvaq 240

Qy 272 LQAWMYKASPEKFTYERRELSAAYIPFGEGDFYHAAIFGGTPTTHILNLTRCPRKGILOD 331  
 Db 241 lqawwykahpdeftyrriksaaylpgqgdfyhaaifggtptqvlntgcfkgllqd 300

Qy 332 KKHIDEAQWHDESHNLKYNKPTKILSPCYCWDYQIGLPSDIKSVKAVQTKYENLVR 391  
 Db 301 kndileaeahdeshlnkyfllnkpskllspsycwdyhgllpsvdfirvklawkkkeynlvr 360

Qy 392 NNV 394  
 Db 361 nnl 363

RESULT 6  
 AAR62508  
 ID AAR62508 standard; Protein; 359 AA.  
 XX AAR62508;  
 AC AAR62508;  
 DT 26-JUN-1995 (first entry)  
 XX Galactosyl transferase clone product.  
 DE Galactosyl transferase clone product.  
 KW Gal-alpha (1,3) galactosyl transferase; xenograft; transplant;  
 KW rejection.  
 XX Sus scrofa domestica.  
 OS WO9421799-A.  
 XX 29-SEP-1994.  
 PD 15-MAR-1994; 94WO-AU00126.  
 PF 16-MAR-1993; 93AU-0007854.  
 PR (AUST-) AUSTIN RES INST.  
 XX McKenzie IFC, Sandrin MS;  
 PI WPI; 1994-317019/39.  
 XX N-PSDB; AAQ74712.  
 DR DNA sequences encoding Gal-alpha (1,3)galactosyl transferase -  
 PT and clones contg. such sequences are used in xenograft therapies  
 XX Disclosure; Page 35; 50pp; English.  
 CC The sequence is that of the product of the porcine Gal-alpha (1,3)  
 CC galactosyl transferase gene which produces a Gal epitope on the  
 CC surface of porcine cells. This epitope is recognised by antibodies  
 CC which are responsible for hyperacute rejection of xenotransplanted  
 CC pig cells, tissues and organs.  
 CC See also AAR62507.  
 XX Sequence 359 AA;

Query Match 71.48; Score 1525; DB 15; Length 359;  
 Best Local Similarity 75.18; Pred. No. 3.3e-145;  
 Matches 271; Conservative 48; Mismatches 38; Indels 4; Gaps 3;

Qy 36 MNVKGKIVLLMLIVSTVVVFWVEYVNRIPVEGENRWOKDMWFPKNGTHSY--OEDNV 93  
 Db 1 mnvkgrrvllmlivstvmvllvlnrnpvegssa-qrgwwfpswfnngthsyheeedai 59



Query Match	71.4%	Score 1525;	DB 17;	Length 359;
Best Local Similarity	75.1%	Pred. NO. 3.3e-145;		
Matches 271;	Conservative 48;	Mismatches 38;	Indels 4;	Gaps
Qy	36	MNVKGVILLMLIVSTVVVVFVWVNRIPVGVGNRQKQHWPFSEFKNGTHISY--QEDNV	93	
Db	1	mnvkgrrvvlsmllvstvmvfvweylnrpevgssa-qrgwfwpswfnngthshyheeedal	59	
Qy	94	EGRRKGRNGRDRIEPEQLMDWFNPKNRPDVLTVTPWKAPIVWEGTYDTALLKYYATQKL	153	
Db	60	gnekeqrkednrgelp-lvdwfnpekrpewvtitrkapvfwegtyrnrvldnyyakql	118	
Qy	154	TVGLTVFAVGKYYIEHYLEDPLESADMYFMVGHRRVIFYVMIDDTSRMPVVLNPLHSLQWF	213	
Db	119	tvgltvfavgryiehyleefllsantcyfmvgkhkviyfmvddlsrmpllielgplrsfkvf	178	
Qy	214	EIRSEKRWQDISMMRWKKTIGEHILAHQHEVDFLCWMDVDVQFDQNFVGVTIGOLVAQLQ	273	
Db	179	eiksekrrwqdismmrmtkigehllahqhevdflcldvdqvfgnrfvgvetlqgsvaqj	238	
Qy	274	AWMYKASPERFTYERRRELSAAYIPFEGEGDFYHAAIFGGTPTTHILNLTRECFKGILODKK	333	
Db	239	awmykahpdeftyerrkesaayipfggdfyhaaifggtptqvlntqecfkgllqdk	298	
Qy	334	HDTEAQWHDSEHLNKYFLFNKPKILSPYECWDYQIGLPSDIKSVKVANOTKEYNLVRNN	393	
Db	299	ndieaeawhdeshlnkyflnkpktllspycwdyhgmsvdlrkvkqwkkeynlvrnn	358	
Qy	394	v	394	
Db	359	i	359	
RESULT	8			
AAW49686				
ID	AAW49686	standard; Protein;	375	AA.
XX				
AC	AAW49686;			
XX				
DT	10-NOV-1998	(first entry)		
XX				
DE	Porcine	alpha-1,3-galactosyl transferase isoform 1.		
XX				
KW	Isoform; porcine; enzyme; alpha-1,3-galactosyl transferase; galactose			
KW	sugar; N-acetylglucosamine; glycoprotein; glycolipid; antibody; pig;			
KW	graft tissue rejection; organ transplantation; xenotransplant.			
XX				
OS	Sus scrofa.			
XX				
PN	FR2751346-A1.			
XX				
PD	23-JAN-1998.			
XX				
PF	19-JUL-1996;	96FR-0009077.		
XX				
PR	19-JUL-1996;	96FR-0009077.		
XX				
XX	(INRM )	INSERM INST NAT SANTE & RECH MEDICALE.		
PA	Pourcel C,	Soullillou JP, Vanhove B;		
PI				
XX				
DR	WPI;	1998-112876/11.		
DR	N-PSDB;	AAV49453.		
XX				
PT	Transgenic non-human donors of organs for human recipients -			
PT	containing DNA encoding antibodies that inhibit graft rejection			
XX				
PS	Claim 4;	Page 32-34;	71pp;	French.
XX				
CC	This sequence represents isoform 1 of the porcine enzyme			
CC	alpha-1,3-galactosyl transferase (alpha-1,3-GT). The enzyme catalyses			



xx Alpha-1,3-galactosyltransferase; alpha-1,3-GalT; transgenic animal;  
KW pig; hyperacute rejection; xenotransplantation; donor organ;  
KW allograft rejection; Gal epitope; gene disruption;  
KW homologous recombination; knock-out.  
XX  
OS Sus scrofa.  
XX W09S20661-A1.  
PN  
XX  
PD  
XX  
XX 03-AUG-1995.  
XX  
PF 27-JAN-1995; 95WO-IB00088.  
XX  
XX 26-JAN-1995; 95US-0188607.  
PR 27-JAN-1994; 94US-0188607.  
XX  
XX (BRES-) BRESATEC LTD.  
PA (SVIN-) ST VINCENT'S HOSPITAL MELBOURNE LTD.  
PA  
XX Crawford RJ, Dapice AJF, Pearse MJ, Rathjen PD;  
PI Robbins AJ;  
XX  
XX WPI: 1995-275446/36.  
DR N-PSDB; AAQ93077.  
XX  
XX New alpha-1,3-galactosyltransferase and leukaemia inhibitor factor  
PT - corresp. DNA and nucleic acid constructs for inactivating the  
PT transferase gene; for eliminating hyperacute region in human  
PT transplants  
XX  
XX Claim 3; Fig.5; 184pp; English.  
PS  
XX  
XX cDNA encoding porcine alpha-1,3-GalT was generated from liver RNA  
CC using primers based on conserved regions of the mouse and cattle alpha-  
CC 1,3-GalT genes. Potential sites to interrupt the alpha-1,3-GalT gene  
CC (via homologous recombination) were identified in exons 4, 7, 8 and 9.  
CC Plugs suitable as donors of organs to overcome hyperacute rejection  
CC problems in human xenotransplantation.  
XX  
XX Sequence 371 AA;  
CC  
Query Match 69.9%; Score 1493; DB 16; Length 371;  
Best Local Similarity 71.8%; Pred. No. 5.9e-142;  
Matches 268; Conservative 49; Mismatches 40; Indels 16; Gaps 4;  
QY 36 MNVKGKVVILLIVSTVVVFWEYVN-----RIPEVGENRWQKDMWFPSPKFN 83  
DB 1 mnvkgrrvllstvmvfweyinspegslfwlqsknpevgssa-qrgwwfpswfn 59  
QY 84 GTHSY--OEDNVGRRKNGRDRIEPEQLWDFNPKNRPDLVTPWKAPIWEGTYDT 141  
DB 60 gthsyheedaagnekeqrkedargelp-lvdwfnpekrpevvltirwkapvwwegtynr 118  
QY 142 ALLEKYYATQKLTGLTFAVGKRYIEHYLEDFLESADMYFMVGHVRVIFVIMDDTSRMPV 201  
DB 119 avldnynakqkicvgtitavrgyiehyieeflisantyfmvghkvifvymvddisrmp 178  
QY 202 VHUNPLHSLOVFIRESKRWQDISMMRMKTIGEHILAHQHEVDFLFCMDVDQVDFDNFG 261  
DB 179 ielgplrsfkvfaketrwqdismmrmktigehilahiheqhevdflfcmddvdqvfqnfg 238  
QY 262 VETLGOLVAQLOAWKYKASPEKTYERRELSAAYIPFGEGDFYHAAIFGCTTHILNLT 321  
DB 239 vecigsvaqlgawwykahnpefeyarrkesaayipfgqgdfyhaaifggtqvtqinit 298  
QY 322 RECFFKGLQDKKHDIEAQWHDESHNLKRYLFNKPRTKILSPEYCDWYQIGLPSDKSVKVA 381  
DB 299 qecfkgilqdkendieaewhdehlnkylilnkptkilspcydcwyhigmsvdirlvkia 358  
QY 382 WQTKYENLVRNNV 394

DB 359 wqkkeynlvrnni 371  
RESULT 11  
AAW49688  
ID AAW49688 standard; Protein; 354 AA.  
XX  
AC AAW49688;  
XX  
DT 10-NOV-1998 (first entry)  
XX  
DE Porcine alpha-1,3-galactosyl transferase isoform 3.  
XX  
KW Isoform; porcine; enzyme; alpha-1,3-galactosyl transferase; galactose;  
KW sugar; N-acetyllactosamine; glycoprotein; glycolipid; antibody; pig;  
KW graft tissue rejection; organ transplantation; xenotransplant.  
XX  
OS Sus scrofa.  
XX FR27511346-A1.  
PN  
XX 23-JAN-1998.  
PD  
XX 19-JUL-1996; 96FR-0009077.  
PF  
XX 19-JUL-1996; 96FR-0009077.  
PR  
XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
PA  
XX Pourcel C, Soullilou JP, Vanhove B;  
PI  
XX WPI: 1998-112876/11.  
DR N-PSDB; AAV49455.  
XX  
XX Transgenic non-human donors of organs for human recipients -  
PT containing DNA encoding antibodies that inhibit graft rejection  
XX  
XX Claim 4; Page 39-41; 71pp; French.  
PS  
XX This sequence represents isoform 3 of the porcine enzyme  
CC alpha-1,3-galactosyl transferase (alpha-1,3-GT). The enzyme catalyses  
CC the attachment of a galactose sugar molecule on the N-acetyllactosamine  
CC moiety found on surface glycoproteins and glycolipids. These sugar  
CC molecules are partly responsible for raising anti-graft antibodies, which  
CC lead to graft tissue rejection. The invention relates to a method of  
CC inhibiting the graft rejection mechanism by introducing the sequence  
CC encoding an antibody targeted to alpha-1,3-GT into the cells of animal,  
CC especially a pig, from whom organs may be used for xenotransplants.  
CC Neutralisation of the alpha-1,3-GT leads to tissues or organs lacking  
CC the galactose on the glycoproteins and glycolipids, thus preventing  
CC induction of the rejection response.  
XX  
XX Sequence 354 AA;  
SQ  
Query Match 67.0%; Score 1431.5; DB 19; Length 354;  
Best Local Similarity 72.2%; Pred. No. 8.7e-136;  
Matches 262; Conservative 48; Mismatches 40; Indels 13; Gaps 5;  
QY 34 KTMNVKGKVVILLIVSTVVVFWEYVNRIPEVGENRWQKDMWFPSPKNGTHSY--QED 91  
DB 3 ktmnvkgrrvllstvmvfweyins-pe-gslfw-----lyqskthsyheeed 52  
QY 92 NVGRRKNGRDRIEPEQLWDFNPKNRPDLVTPWKAPIWEGTYDTALLEKYYATQ 151  
DB 53 aignekeqrkedargelp-lvdwfnpekrpevvltirwkapvwwegtynravldnyyak 111  
QY 152 KLTGVLTFAVGKRYIEHYLEDFLESADMYFMVGHVRVIFVIMDDTSRMPVHNLPLHSLO 211  
DB 112 kicvgtitavrgyiehyieeflisantyfmvghkvifvymvddisrmpielgplrsfk 171  
QY 212 VFEIRSEKRWQDISMMRMKTIGEHILAHQHEVDFLFCMDVDQVDFDNFGVETLGOLVAQ 271

```

Db 172 vfeiksekrwgdismmrkktigeihlahiqhevdflfcmddvqvgfnfgvetlgsqvaq 231
QY 272 LOAWMYKASPEKFTYERRELSSAAVIFPGEGDYFYHAAIFGCTPTHILNLTRECFKGILOD 331
Db 232 lqawwykahpdefcyerrkesaayipfgqgdfyyhaaifggtqvtqiniqecfkgllqd 291
QY 332 KKHIDIAQWHDHSHLNKYFLFNKPTKILSPCYCWDYQIGLPSDIKSVKVAQTKKEYNLVR 391
Db 292 kendieaewhdeshlnkyflfnkptkilspeycwdyhgmsvdlrvkvwkqkkeynlvr 351
QY 392 NNV 394
Db 352 nni 354

RESULT 12
AAW49689
ID AAW49689 standard; Protein; 342 AA.
XX
AC AAW49689;
XX
DT 10-NOV-1998 (first entry)
XX
DE Porcine alpha-1,3-galactosyl transferase isoform 4.
XX
KW Isoform; porcine; enzyme; alpha-1,3-galactosyl transferase; galactose;
KW sugar; N-acetyllactosamine; glycoprotein; glycolipid; antibody; pig;
KW graft tissue rejection; organ transplantation; xenotransplant.
XX
OS Sus scrofa.
XX
PN FR2751346-A1.
XX
PD 23-JAN-1998.
XX
PF 19-JUL-1996; 96FR-0009077.
XX
PR 19-JUL-1996; 96FR-0009077.
XX
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Pourcel C, Soullilou JP, Vanhove B;
XX
DR WPI; 1998-112876/11.
XX
DR N-PSDB; AAV49456.
XX
PT Transgenic non-human donors of organs for human recipients -
PT containing DNA encoding antibodies that inhibit graft rejection
XX
PS Claim 4; Page 43-44; 71pp; French.
XX
CC This sequence represents isoform 4 of the porcine enzyme
CC alpha-1,3-galactosyl transferase (alpha-1,3-GT). The enzyme catalyses
CC the attachment of a galactose sugar molecule on the N-acetyllactosamine
CC moiety found on surface glycoproteins and glycolipids. These sugar
CC molecules are partly responsible for raising anti-graft antibodies, which
CC lead to graft tissue rejection. The invention relates to a method of
CC inhibiting the graft rejection mechanism by introducing the sequence
CC encoding an antibody targeted to alpha-1,3-GT into the cells of animal,
CC especially a pig, from whom organs may be used for xenotransplants.
CC Neutralisation of the alpha-1,3-GT leads to tissues or organs lacking
CC the galactose on the glycoproteins and glycolipids, thus preventing
CC induction of the rejection response.
XX
SQ Sequence 342 AA;

Query Match 66.9%; Score 1428.5; DB 19; Length 342;
Best Local Similarity 71.3%; Pred. No. 1.7e-135;
Matches 259; Conservative 45; Mismatches 34; Indels 25; Gaps 3;

QY 34 KIMNVKGVILLMLIVSTVVWFVWEYVNRIPVEGENRWQDMWFPSPFKNGTHSY--QED 91

```

```

Db 3 kinnvkgrrvvlsmllivstvmvfweylnr-----thsyheeed 40
QY 92 NVEGRRKKGNGDRIRIEPQOLWDNPNKRNPDVLTVTWKAPIVWEGTYDTALLKYYATO 151
Db 41 aignekeqrkednrgelp-lvdwfnpekrpevvtirwkapvwwegtyrnrvlndnyyakq 99
QY 152 KLTVGLTFVAVGKYIEHYLEDLFESADMYPMVGHVRVIFYVMIDDTSRMPVHNLPLHSLQ 211
Db 100 kitvgltvavgrylehyleeftlsantyemvghkvifylmvddlsrmpfllieqlrfsfk 159
QY 212 VFEIRSEKRWQDISMRMKTIGEHLAHIQHEVDVDFLCMDVDQVQFQDNFGVETLGLQVAQ 271
Db 160 vfeiksekrwgdismmrkktigeihlahiqhevdflfcmddvqvgfnfgvetlgsqvaq 219
QY 272 LOAWMYKASPEKFTYERRELSSAAVIFPGEGDYFYHAAIFGCTPTHILNLTRECFKGILOD 331
Db 220 lqawwykahpdefcyerrkesaayipfgqgdfyyhaaifggtqvtqiniqecfkgllqd 279
QY 332 KKHIDIAQWHDHSHLNKYFLFNKPTKILSPCYCWDYQIGLPSDIKSVKVAQTKKEYNLVR 391
Db 280 kendieaewhdeshlnkyflfnkptkilspeycwdyhgmsvdlrvkvwkqkkeynlvr 339
QY 392 NNV 394
Db 340 nni 342

RESULT 13
AAR62507
ID AAR62507 standard; Protein; 313 AA.
XX
AC AAR62507;
XX
DT 26-JUN-1995 (first entry)
XX
DE Galactosyl transferase 3' clone product.
XX
KW Gal-alpha (1,3) galactosyl transferase; xenograft; transplant;
KW rejection.
XX
OS Sus scrofa domestica.
XX
PN WO9421799-A.
XX
PD 29-SEP-1994.
XX
PF 15-MAR-1994; 94WO-AU00126.
XX
PR 16-MAR-1993; 93AU-0007854.
XX
PA (AUST-) AUSTIN RES INST.
XX
PI McKenzie IFC, Sandrin MS;
XX
DR WPI; 1994-317019/39.
DR N-PSDB; AAQ74711.
XX
PT DNA sequences encoding Gal-alpha (1,3)galactosyl transferase -
PT and clones contg. such sequences are used in xenograft therapies
XX
PS Disclosure; Page 31; 50pp; English.
XX
CC The sequence is that of the product of the porcine Gal-alpha (1,3)
CC galactosyl transferase gene which produces a Gal epitope on the
CC surface of porcine cells. This epitope is recognised by antibodies
CC which are responsible for hyperacute rejection of xenotransplanted
CC pig cells, tissues and organs.
CC See also AAR62508.
XX
SQ Sequence 313 AA;

```



